BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2017-55

Risk-Neutral Moment-Based Estimation of Affine Option Pricing Models

by Bruno Feunou and Cédric Okou

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2017-55

December 2017

Risk-Neutral Moment-Based Estimation of Affine Option Pricing Models

by

Bruno Feunou[1] and Cédric Okou[2]

[1] Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
feun@bankofcanada.ca

[2] ESG
UQAM
okou.cedric@uqam.ca

ISSN 1701-9397

© 2017 Bank of Canada

Acknowledgements

An earlier version of this paper was circulated and presented at various seminars and conferences under the title "Affine Term Structure of Risk-Neutral Moments Models." We thank seminar participants at the University of Western Ontario for helpful comments. We are grateful for fruitful conversations with Jean-Paul Renne, Guillaume Rousselet, and Nicola Fusari. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada.

Abstract

This paper provides a novel methodology for estimating option pricing models based on risk-neutral moments. We synthesize the distribution extracted from a panel of option prices and exploit linear relationships between risk-neutral cumulants and latent factors within the continuous time affine stochastic volatility framework. We find that fitting the Andersen, Fusari, and Todorov (2015b) option valuation model to risk-neutral moments captures the bulk of the information in option prices. Our estimation strategy is effective, easy to implement, and robust, as it allows for a direct linear filtering of the latent factors and a quasi-maximum likelihood estimation of model parameters. From a practical perspective, employing risk-neutral moments instead of option prices also helps circumvent several sources of numerical errors and substantially lessens the computational burden inherent in working with a large panel of option contracts.

Bank topics: Asset pricing; Econometric and statistical methods
JEL code: G12

Résumé

Dans cette étude, nous proposons une méthodologie nouvelle pour estimer les modèles d'évaluation d'options, fondée sur les moments neutres à l'égard du risque. Nous synthétisons la distribution extraite de notre échantillon de prix d'options et exploitons les relations linéaires qui existent entre les cumulants neutres à l'égard du risque et les variables latentes dans le cadre d'un modèle affine à volatilité stochastique en temps continu. Nous établissons que l'ajustement du modèle d'évaluation d'options d'Andersen, Fusari et Todorov (2015b) aux moments neutres à l'égard du risque permet de saisir le gros de l'information véhiculée par les prix des options. Nous jugeons notre stratégie d'estimation à la fois efficace, facile à mettre en œuvre et robuste, puisqu'elle autorise un filtrage linéaire direct des variables latentes et une estimation des paramètres des modèles par la méthode du quasi-maximum de vraisemblance. Du point de vue pratique, le recours aux moments neutres à l'égard du risque, au lieu des prix des options, permet d'éviter plusieurs sources d'erreurs numériques et réduit substantiellement la masse de calculs à effectuer pour traiter un vaste échantillon d'options.

Sujets : Évaluation des actifs ; Méthodes économétriques et statistiques
Code JEL : G12

Non-Technical Summary

Option prices are routinely used to infer the distribution of future movements in their underlying asset prices. For example, financial econometricians use option prices to precisely evaluate the variance, the skewness, and the kurtosis of the aforementioned distribution.

However, academics and practitioners do not use such sophisticated option pricing models because of estimation challenges. These hurdles are (1) the computational burden inherent in working with a large panel of option contracts, (2) the numerical approximations of ordinary differential equations to compute the characteristic function, (3) the numerical integrations of the characteristic function to compute option prices, and (4) the nonlinear filtering of the latent state variables from option prices.

This paper proposes to estimate these models using the forward-looking variance, skewness, and kurtosis instead of raw option prices data. These forward-looking moments can be thought of as portfolios of options contracts, along the strike dimensions. The methodology used to compute them is similar to the VIX methodology and is discussed extensively in the literature.

Our proposal is effective, easy to implement, and robust, because it allows us to infer not only model parameters, but unobserved factors as well. From a practical perspective, employing the forward-looking variance, skewness, and kurtosis instead of option prices also helps circumvent several sources of numerical errors and substantially lessens the computational burden inherent in working with a large panel of option contracts.

1 Introduction

Option prices are of importance to investment decisions, as they provides useful forward-looking information on market conditions. Thus, a large body of the asset pricing literature aims at designing valuation models that can accurately fit the observed option prices. Most state-of-the-art option pricing specifications account for the salient empirical regularities of the underlying distribution – such as volatility randomness and persistence, as well as substantial conditional tail thickness – in the specification of stochastic volatility dynamics and (time-varying intensity) jump components. However, the implementation and estimation of continuous time jump-diffusion models is challenging because of the econometric and computational complexity involved when using available option prices over long periods of time. Simply put, the main challenge in stochastic volatility model estimation is that latent state variables must be inferred along with model parameters from asset prices. Keep in mind that, as Bates (2006) points out, the relationship between option prices and latent state variables is highly nonlinear. One may argue that the solution to this problem is simple. One could ignore some data, for instance by focusing on a subset of contracts (at-the-money or Wednesday options). Clearly, this choice is inefficient from a statistical perspective. Alternatively, one could re-estimate the model using shorter samples. In the limit, one could re-estimate the model every day, as in Bakshi et al. (1997). Because these models are genuinely dynamic, this approach is also not optimal. The estimation challenges of stochastic volatility models might explain why generalized autoregressive conditional heteroskedasticity (GARCH)-type models remain very popular in option pricing.

This article proposes a new and generic estimation approach that uncovers the latent state variables (unobserved factors) by synthesizing the risk-neutral distribution extracted from a large panel of option prices spanning several dates, maturities, and moneyness. Bakshi and Madan (2000) and Bakshi et al. (2003) discuss how to build conditional risk-neutral moments from option prices without resorting to any parametric assumption. Model-free risk-neutral moments can be used to identify latent factors, and thus circumvent one of the major challenges in estimating stochastic volatility models. Indeed, using option prices for stochastic volatility model estimation usually involves Fourier transform inversions that entail high-dimensional numerical integrations, especially in the presence of several unknown model parameters and latent factors. To estimate continuous time affine-stochastic volatility models, we use risk-neutral moments (instead of option prices) and exploit the linear function linking any conditional cumulant to the latent factors within the affine

family.[1] The affine relationship between cumulants and factors enables us to extract the unobserved state variables and estimate the model parameters using a slightly modified version of the linear Kalman filter. Let us stress that it is the cumulants (not the moments) that are linearly related to the latent factors in affine models. Interestingly, cumulants can be easily expressed in terms of moments. Therefore, we will use the terms "cumulants" and "moments" whenever is appropriate.[2] Ultimately, we can undo the nonlinearities between option prices and factors by transforming option prices into risk-neutral cumulants. Moreover, employing option-implied cumulants instead of option prices mitigates the difficulty inherent in working with a large number of option contracts for the estimation. Namely, instead of using 570,108 option contracts, we "aggregate" them into 73,752 risk-neutral cumulants. Our approach bears some resemblance to the estimation of affine term structure of interest rates models, where bond yields are linearly related to the unobserved factors. A thorough review of the affine term structure framework can be found in Singleton (2006), and Joslin et al. (2011).

Our main contribution is to establish that, for the estimation of option pricing models, fitting risk-neutral (second, third, and fourth) cumulants overcomes the nonlinearities and subsumes a sizeable fraction of the information content of option prices. Using the Andersen et al. (2015b) (henceforth, AFT) model, we show that the risk-neutral moment-based estimation of stochastic jump-diffusion specifications within the affine-Q family is effective, simple, and provides robust results. As in the case of most state-of-the-art multi-factor pricing models, the ordinary differential equations satisfied by the factor loadings (functionals of the model's primitive parameters) of the conditional characteristic function do not admit explicit analytical solutions. Interestingly, we have computed their derivatives (cumulants) in explicit closed-form. This observation further underscores a key implementation advantage that our risk-neutral moment-based estimation approach offers.

Our work is related to a growing literature that relies on various observable quantities to fit continuous-time option pricing models. Recent studies by Duan and Yeh (2010) and Kaek and Alexander (2012), use the VIX index to uncover the unobserved variance path before proceeding to the estimation of option pricing models. However, in the presence of jumps in the underlying return

[1] The stochastic volatility family considered is affine under the risk-neutral probability distribution. This assumption is needed to derive the analytical expressions of risk-neutral cumulants in closed-form. Nonetheless, this assumption is not very restrictive because the corresponding processes may be non-affine under the physical probability distribution.

[2] The empirical behavior of option-implied moments (rather than cumulants) have been extensively studied in the literature on derivatives.

process, the squared VIX index is a biased proxy for the risk-neutral expectation of the quadratic variation of log returns: see, e.g., the discussion in Carr et al. (2012). As in Bakshi and Madan (2000) and Chang et al. (2012), our construction of nonparametric risk-neutral moments minimizes jump approximation and discretization errors. Another line of research fits option models after extracting the latent spot variance from observed variance swap rates, as implemented in Egloff et al. (2010), Amengual and Xiu (2015), and Aït-Sahalia et al. (2015). These studies focus on fitting variance swap prices, but do not assess the option pricing accuracy. Instead, we evaluate the option fitting performance. In addition and perhaps more importantly, we go beyond the second-order risk-neutral moment and fit model-implied to observed (nonparametric) third- and fourth-order risk-neutral moments at different maturities.

Regarding the estimation strategies of stochastic volatility models, several approaches have been proposed. Following Bakshi et al. (1997), Bates (2000), and Huang and Wu (2004) use a two-step scheme that filters the unobserved factors and estimates the structural parameters, iteratively. Christoffersen et al. (2009), and Johannes et al. (2009) implement improved variants of this particle-filtering algorithm for general specifications of jump-diffusion processes. A similar estimation methodology is adopted by Jones (2003) and Eraker (2004) within a Bayesian framework. Carr and Wu (2007) opt for a Kalman filter approach. Alternatively, stochastic volatility models can be estimated by the efficient method of moments, as pointed out by Gallant and Tauchen (1996), Andersen et al. (2002), and Chernov and Ghysels (2000). Gagliardini et al. (2011) develop a test strategy based on extended method of moments. Pan (2002) runs a generalized method of moment estimation, while Bates (2006) uses a maximum likelihood methodology to fit latent affine processes. Feunou and Tédongap (2012) employ a GARCH approximation of stochastic volatility dynamics. In a recent study, Andersen et al. (2015a) add a penalization term to Bates' (2000) objective function to estimate risk-neutral parameters and latent factors. This penalization term captures the gap between observed and model-implied spot variance, and is designed to discipline the estimated factors. Moreover, Andersen et al. (2015a) provide a strong theoretical framework for conducting inference. Within the affine framework, although option prices are nonlinear function of factors, there exist portfolios of these options contracts weighted across moneyness (risk-neutral moments) that enable us to undo the nonlinearity and apply a modified linear Kalman filtering technique for the estimation. Note that the objective of our study is not to propose another option pricing model. We focus on providing a new methodology for estimating affine option pricing models by fitting the observed risk-neutral cumulants. To the best of our knowledge, we are the

first to investigate the joint fitting of observed higher-order (second, third, and fourth) risk-neutral cumulants across different maturities.

The remainder of the paper is organized as follows. Section 2 presents the general affine-Q class of stochastic volatility models, reviews the estimation challenges, and outlines the core concept of the risk-neutral moment-based methodology. Section 3 discusses the proposed risk-neutral moment-based estimation strategy. Section 4 implements the risk-neutral moment-based estimation approach for the AFT model. Section 5 describes the data and the computation of nonparametric risk-neutral quantities, contrasts the performance of the moment-based estimation method with that of the AFT estimation approach in delivering realistic option prices and risk-neutral quantities, and assesses the implied risk premia. Section 6 concludes.

2 General affine-Q framework

Consider the time t price S_t of a security, and define the log-price process as $\{y_t = \log(S_t)\}_{t\geq 0}$. The information available up to the current time t is characterized by a set of progressive filters $(\mathcal{F}_t)_{t\geq 0}$. Let F be an $N \times 1$ vector of factors governing the distribution of y. We posit that the discount rate function is affine in the factors, and given by

$$R_t = r_f + \rho_r' F_t,$$

where r_f is the log risk-free rate and ρ_r is an $N \times 1$ vector of factor loadings.

For a future payoff date $T \geq t$ and $u \in \mathbb{C}$, models under consideration are those for which the following transform function

$$\psi(u; y_t, F_t, T, r_f, \rho_r) \equiv E_t^Q \left[\exp\left(-\int_t^T R_s ds \right) e^{uy_T} \right]$$

exists and is exponential affine in F, with E_t^Q denoting the risk-neutral expectation conditioned on $\mathcal{F}_t$. Following Duffie et al. (2000) and Duffie et al. (2003), we assume that all technical regularity conditions are satisfied so that the expectation is well defined.[3]

[3] This set of assumptions is discussed on page 1351 in Duffie et al. (2000).

By setting $\tau = T - t$, the conditional characteristic function takes the form[4]

$$\psi(u; y_t, F_t, T, r_f, \rho_r) = \exp\left(u y_t + \alpha(u; \tau, r_f, \rho_r) + \beta(u; \tau, r_f, \rho_r)' F_t\right), \tag{1}$$

where, in continuous time, $\alpha(\bullet)$ and $\beta(\bullet)$ are solutions to ordinary differential equations (ODEs). We refer the reader to Duffie et al. (2000) and Duffie et al. (2003) for general expressions of these ODEs.[5]

In some special cases, the explicit solutions to these ODEs exist. Thus, the analytical expressions of $\alpha(\bullet)$ and $\beta(\bullet)$ are known in closed-form, as in Heston (1993), and Huang and Wu (2004). In other realistic cases, the ODEs must be solved numerically, in particular for state-of-the-art models featuring self-exciting jumps as in the AFT model, among others. Even though a few rich option pricing specifications have been recently proposed in the literature, our empirical study focuses on the AFT model for the sake of conciseness. The three-factor AFT model provides a realistic framework for option contract valuation.

2.1 Standard option pricing and estimation

2.1.1 Option valuation

The extant option pricing literature relies on contract values to back out model parameters and factors. For affine models, the general pricing formula of a security that pays e^{ay_T} at time $T = t + \tau$ when the event $by_T \leq x$ occurs, with any $a, b \in \mathbb{R}$, is given by (see Proposition 2.2 in Duffie et al. (2000))

$$\begin{aligned} G_{a,b}(x, y_t, F_t, T) &= E_t^Q\left[\exp\left(-\int_t^T R_s ds\right) e^{ay_T} \mathbb{1}_{by_T \leq x}\right], \\ &= \frac{\psi(a; y_t, F_t, T, r_f, \rho_r)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\Im\left[\psi(a + i\nu b; y_t, F_t, T, r_f, \rho_r) e^{-iwx}\right]}{\nu} d\nu, \end{aligned} \tag{2}$$

where $\Im(u)$ returns the imaginary part of $u \in \mathbb{C}$. Consequently, a European call option with payoff $(e^{y_T} - e^x)^+$ is priced (see Equation 1.6 in Duffie et al. (2000)) at

$$C(x, y_t, F_t, \tau) = G_{1,-1}(-x, y_t, F_t, T) - e^x G_{0,-1}(-x, y_t, F_t, T), \tag{3}$$

[4] Formally, ψ gives the conditional characteristic function when the discount rate $R_t = 0$.

[5] In discrete time, $\alpha(\bullet)$ and $\beta(\bullet)$ are solutions to recursive finite difference equations, as studied in Darolles et al. (2006).

where $X = e^x$ gives the strike price. A few important remarks arise from the pricing rules in Equations 2 and 3. Namely, there are at least three potential sources of numerical errors that stand as challenges from a practical implementation standpoint.

First, both the integral and integrand (ψ) in the pricing relation must be approximated numerically. This entails a high-dimensional numerical integration, especially when there is a large number of factors. The numerical challenge becomes more critical in estimation settings, where parameter values are not fixed but must be inferred along with factor estimates.

Second, the highly nonlinear link between option prices and factors precludes the use of standard linear filters and often requires complex filtering techniques, such as the square-root unscented Kalman filter (Van der Merwe and Wan (2001)).

Third, as mentioned in the introduction, the sheer size of a typical option price panel adds to the complexity of option pricing model estimation. The observed values of option contracts written on a given underlying asset form a panel along the time dimension (observation date), the maturity dimension (tenor), and the cross-section dimension (moneyness). The option panel in this study contains a large number of records (570,108) along those three dimensions. Note that to circumvent the computational burden, some estimation procedures select a subset of options, such as at-the-money or Wednesday contracts.

2.1.2 Andersen et al. (2015a) estimation

We outline the Andersen et al. (2015a) estimation procedure that is used as a reference point for the proposed risk-neutral moment-based estimation strategy. On a given day $t = 1, \ldots, T$ in a typical option panel, we observe N_t contract values for various strike prices and maturities. As is standard in the literature on derivatives, we rely on the vega-weighted root mean squared error

$$VWRMSE_t \equiv \sqrt{\frac{1}{N_t}\sum_{j=1}^{N_t}\left(\left(C_j^{Mkt} - C_j^{Mod}\right)/BSV_j^{Mkt}\right)^2},$$

where C_j^{Mkt} is the j^{th} option contract value observed on the market, C_j^{Mod} is the corresponding model-implied price, and BSV_j^{Mkt} is the observed Black-Scholes vega of the option. Note that the VWRMSE is a computationally cheaper alternative to the implied volatility root mean squared error.[6]

[6] The implied volatility root mean squared error $IVRMSE_t \equiv \sqrt{\frac{1}{N_t}\sum_{j=1}^{N_t}\left(IV_j^{Mkt} - IV_j^{Mod}\right)^2}$ measures the wedge between the observed or market-based implied volatility $IV_j^{Mkt} = BS^{-1}\left(C_j^{Mkt}\right)$ and the model-based implied volatil-

The optimization is performed according to

$$\left(\{\hat{F}_t\}_{t=1,\ldots,T}, \hat{\theta}\right) = \underset{\{F_t\}_{t=1,\ldots,T},\theta}{\arg\min} \sum_{t=1}^{T} \left\{ VWRMSE_t^2 + \omega_m \left(\sqrt{V_t^m} - \sqrt{V_t}\right)^2 \right\}, \tag{4}$$

where ω_m is a nonnegative penalization coefficient, V_t is the spot volatility, and V_t^m is a nonparametric high-frequency estimator of spot volatility computed using a fine grid of n records within a unit interval of time. While the spot volatility V_t is affine the factors, the pricing error in the VWRMSE is not. The objective function in Equation 4 incorporates a penalizing term to ensure that the model-implied spot volatility aligns well with its model-free high-frequency estimate. Following AFT, we set $\omega_m = 0.05$ and construct a consistent estimator of the spot variance at the end of each trading day, using a one-minute grid of the underlying returns, as

$$V_t^m = \frac{n}{m_n} \sum_{i=n-m_n+1}^{n} \left(y_{t+i/n} - y_{t+(i-1)/n}\right)^2 \mathbb{1}\left(\left|y_{t+i/n} - y_{t+(i-1)/n}\right| \leq \delta n^{-\varpi}\right). \tag{5}$$

Over 6.5 hours in a typical trading day, a one-minute grid contains about $n = 390$ observations, and the value of m_n corresponds to a fraction (3/4) of n. The remaining tuning parameters are set as follows: $\varpi = 0.49$ and $\delta = \sqrt{BV_{t-1} \wedge RV_{t-1}}$, where BV_{t-1} and RV_{t-1} are high-frequency bi-power variation and realized volatility estimators.

As we deal with a large panel of option data (570,108 contracts recorded over 3,073 days), we implement an iterative two-step procedure for the AFT estimation. Namely, parameters and unobservable latent factors are sequentially estimated. In the first step, for a given set of structural parameters, and on each observation date t, we minimize the pricing errors (the term in the curly brackets in Equation 4) to get estimates of latent factors. In the second step, given the set of latent factors estimated from the first step for all dates $t = 1, ..., T$, we solve one aggregate sum of squared pricing errors optimization problem to get a new estimation of the model parameters. The procedure iterates between these two steps until the reduction in the overall objective in the second step becomes marginal (convergence). The outlined challenges motivate our alternative estimation methodology, which synthesizes 570,108 raw option prices into 73,752 risk-neutral cumulants, and use them to fit affine option models.

ity $IV_j^{Mod} = BS^{-1}(C_j^{Mod})$, where BS^{-1} stands for the inverse of the Black-Scholes formula. Renault (1997) argues for using the IVRMSE as a performance metric in the appraisal of option pricing models. However, given the large number of contracts (570,108) that we consider, a direct computation of the IVRMSE is costly because each option must be inverted to obtain the corresponding implied volatility. To circumvent this computational challenge, we opt for the VWRMSE.

2.2 Our approach: pricing with a portfolio of options

We conjecture that the risk-neutral variance, skewness, and kurtosis summarize most of the information embedded in option prices. Specifically, risk-neutral moments are weighted portfolios of options. In fact, any twice-continuously differentiable payoff with bounded expectation can be spanned (Bakshi and Madan (2000)) according to the formula

$$\begin{aligned} \mathcal{G}[S] \quad &= \mathcal{G}[\overline{S}] + (S-\overline{S})\mathcal{G}_S[\overline{S}] + \int_{\overline{S}}^{\infty} \mathcal{G}_{SS}[X](S-X)^+ dX \\ &+ \int_0^{\overline{S}} \mathcal{G}_{SS}[X](X-S)^+ dX, \end{aligned} \tag{6}$$

which corresponds to positions in the slope (first derivative $\mathcal{G}_S[\bullet]$ evaluated at some $\overline{S}$) and the curvature (second derivative $\mathcal{G}_{SS}[\bullet]$ evaluated at the strike price X) of the payoff function. Thus, the price of a contingent claim is computed as

$$\begin{aligned} E_t^Q\{e^{-r\tau}\mathcal{G}[S]\} \quad &= e^{-r_{t,\tau}}(\mathcal{G}[\overline{S}] - \overline{S}\mathcal{G}_S[\overline{S}]) + \mathcal{G}_S[\overline{S}]S_t + \int_{\overline{S}}^{\infty} \mathcal{G}_{SS}[X]C(t,\tau;X)dX \\ &+ \int_0^{\overline{S}} \mathcal{G}_{SS}[X]P(t,\tau;X)dX, \end{aligned} \tag{7}$$

reflecting the value of a weighted portfolio that includes risk-free bonds, the underlying asset, and out-of-the-money (OTM) calls and puts. The log-return on the underlying asset value ($S = e^y$) between time t and $t+\tau$ is denoted by $r_{t,\tau} = y_{t+\tau} - y_t$. Accordingly, higher-order model-free or observed risk-neutral moments are constructed with a payoff function $\mathcal{G}[S] = r_{t,\tau}^n$ describing power ($n = 2$-quadratic, 3-cubic, and 4-quartic) contracts.[7] Given that cumulants are related to the expected payoff of power contracts $\chi_n \equiv E_t^Q\left[r_{t,\tau}^n\right]$, with

$$\begin{cases} CU M_{t,\tau}^{(2)Q} = \chi_2 - \chi_1^2, \\ CU M_{t,\tau}^{(3)Q} = \chi_3 - 3\chi_2\chi_1 + 2\chi_1^3, \\ CU M_{t,\tau}^{(4)Q} = \chi_4 - 4\chi_3\chi_1 - 3\chi_2^2 + 12\chi_2\chi_1^2 - 6\chi_1^4, \end{cases} \tag{8}$$

they can be assessed in the data using weighted portfolios of OTM options, as in Equation 7.

Moreover, the model-implied formulas for all existing model-implied risk-neutral moments – expressed as functions of the model's primitive parameters – can be obtained by deriving the

[7] An important caveat is that the moment-replicating portfolios require an infinite range of strikes to be observed, and hence, truncation of this infinite range of strikes can lead to serious biases in the estimates of risk-neutral moments. Moreover, long-maturity risk-neutral moments require the use of illiquid long-dated options.

characteristic function $\psi_t(u;T,0,0)$. The n^{th} derivative of the log characteristic function (also known as cumulant) is computed as

$$\begin{aligned} CUM_{t,\tau}^{(n)Q} &\equiv \left.\frac{\partial^n \ln \psi_t(u;T,0,0)}{(\partial u)^n}\right|_{u=0}, \\ &= \left.\frac{\partial^n \alpha(u;\tau,0,0)}{(\partial u)^n}\right|_{u=0} + \left.\frac{\partial^n \beta(u;\tau,0,0)}{(\partial u)^n}\right|_{u=0} F_t, \\ &\equiv A_\tau^{(n)} + B_\tau^{(n)} F_t. \end{aligned} \tag{9}$$

For $n \in \mathcal{O} = \{2,3,4\}$, the second (resp. third and fourth) cumulant corresponds to the second (resp. third and fourth) derivative of the log characteristic function with respect to u, evaluated at $u = 0$.

The linear relationship in Equation 9 is central to our estimation methodology. It offers an analytical mapping between any given observed n^{th}-order cumulant and latent factors, weighted by functionals of the model's primitive parameters. To estimate affine-Q models, we exploit Equation 9 instead of the option pricing Equation 2 of Duffie et al. (2000). Our approach is appealing, because it avoids the costly numerical task of approximating high-dimensional integrals. It also helps circumvent the challenge of working with a large option panel, as prices are "aggregated" in informative portfolios (or risk-neutral moments). Furthermore, by taking advantage of the linear link between cumulants and factors, a straightforward linear filtering technique can be applied to obtain parameter and factor estimates.

As alluded to above, $A_\tau^{(n)}$ and $B_\tau^{(n)}$ in Equation 9 are functions of model parameters, cumulant orders (n), and maturities (τ). To provide a precise illustration of these functionals, let us consider the widely used Heston (1993) model. This single-factor model is given by

$$\begin{aligned} \frac{dS_t}{S_t} &= r dt + \sqrt{V_t} dW_t^Q, \\ dV_t &= (a - bV_t)\, dt + \sigma \sqrt{V_t} dB_t^Q, \end{aligned}$$

where $\left(W_t^Q, B_t^Q\right)$ is a two-dimensional Brownian motion with $corr\left(W_t^Q, B_t^Q\right) = \rho$.

Appendix A gives the explicit expressions of the location $A_\tau^{(n)}$ and slope $B_\tau^{(n)}$ coefficients in the factor representation of cumulants. It is worth mentioning that these formulas not only link cumulants to model parameters and factors, but also highlight the importance of higher-order cumulants in identifying key model parameters. For instance, the leverage effect is known to drive

the asymmetry as well as the fat-tailedness of the conditional return distribution. To illustrate the distinct implications of leverage, we plot the factor location and slope coefficients of each cumulant against ρ, using a fine grid of values ranging from -1 to 1. As in Heston (1993), the remaining model parameter values are set to $a = 0.02$, $b = 2$, $\sigma = 0.1$, and $\tau = 0.5$. Figure 1 shows that the leverage effect ($\rho < 0$) linearly increases the factor loading for the variance, whereas it decreases that of the third cumulant. For the fourth cumulant, we see a parabolic increase in the factor coefficients as ρ moves away from 0 toward -1 or 1. These observations are consistent with well-established empirical facts and suggest that, beyond the variance, the third and fourth cumulants contain relevant information that is useful in estimating the leverage effect with actual data.

We now present our estimation methodology that fits the parameters and the factors of affine-Q models using risk-neutral moments.

3 Estimation and inference

3.1 The affine-Q moment-based estimation strategy

We focus on the estimation of affine jump-diffusion pricing models. Our approach borrows from the affine term structure of the interest rate literature. The intuition is simple: since cumulants are linear in the factors, we can use them as observed quantities to pin down model parameters and reveal unobserved factors. Thus, we circumvent a major challenge in estimating latent factor models.

Given that our framework is affine, the linear Kalman filter appears as a natural estimation technique. The affine-Q models can be easily cast in a (linear) state-space form where the measurement equations relate the observed or model-free risk-neutral cumulants to the latent factors (state variables), and the transition equations describe the dynamic of these factors.

Assume that on each day we observe risk-neutral cumulants computed at J different maturities. These risk-neutral cumulants are linearly linked to the state vector within the affine jump-diffusion family. Obviously, the number of n^{th}-order risk-neutral cumulants observed on a given day is often far greater than the dimension of the latent factor ($J >> N$). For instance, in the Heston (1993) model, the state vector simply contains the stochastic diffusion ($N = 1$). For a given day t, let's stack together the n^{th}-order risk-neutral cumulant observed at distinct maturities in a vector denoted by $CUM_t^{(n)Q} = (CUM_{t,\tau_1}^{(n)Q}, ..., CUM_{t,\tau_J}^{(n)Q})'$, where $n \in \mathcal{O} = \{2, 3, 4\}$

and $card(\mathcal{O}) = 3$. Let's further stack the second, third, and fourth cumulant vectors in $CUM_t^Q = (CUM_t^{(2)Q}, CUM_t^{(3)Q}, CUM_t^{(4)Q})'$ to build a $3J \times 1$ vector. This implies the following measurement equation

$$CUM_t^Q = \Gamma_0 + \Gamma_1 F_t + \Omega^{1/2} \vartheta_t, \tag{10}$$

where the dimension of the unobserved state vector F_t is $N \times 1$. Notably, Γ_0 and Γ_1 are $3J \times 1$ and $3J \times N$ matrices of coefficients, whose analytical expressions depend on $A^{(u)}$ and $B^{(u)}$ in Equation 9. The last term in Equation 10 is a vector of observation errors, where Ω is a $3J \times 3J$ diagonal covariance matrix, and ϑ_t denotes a $N \times 1$ vector of independent and identically distributed (i.i.d.) standard Gaussian disturbances. The state propagation relation comes from the discretization of the Euler equation and writes

$$F_{t+1} = \Phi_0 + \Phi_1 F_t + \Sigma(F_t)^{1/2} \xi_{t+1}, \tag{11}$$

where $\Sigma(F_t) \equiv Var_t(F_{t+1})$, the conditional covariance matrix of F_{t+1}, is a known affine function of F_t. In the state dynamics, Φ_0 and Φ_1 are $N \times 1$ and $N \times N$ matrices of coefficients. The transition equation noise ξ_{t+1} is a $N \times 1$ zero-mean normalized martingale difference vector that is assumed to be independent from ϑ_t.

It is important to stress that, while the measurement relation in Equation 10 reflects the risk-neutral distribution (and thus depends exclusively on risk-neutral parameters), the state vector dynamics in Equation 11 should be specified under the physical probability measure. Therefore, some parameters should be shifted to reflect risk premia.[8] Thus, our state-space representation allows for the identification of risk-neutral and risk-premium parameters. In this regard, our risk-neutral moment-based estimation approach differs fundamentally from those implemented by Bates (2000) and by AFT, which only identify risk-neutral parameters. Nonetheless, since the risk-neutral parameters estimation is the focus of this paper, we assume zero risk-premium from now on, and defer risk-premium parameters estimation to Section 5.3.

Note that the Kalman filter is not optimal in this case, given that the conditional covariance $Var_t(F_{t+1})$ depends on the latent vector F_t, and thus, is intrinsically unknown. To circumvent this challenge, we employ a slightly modified version of the standard Kalman filter, where $\Sigma(F_{t|t})$ is used as an estimate of $\Sigma(F_t)$ at iteration $t+1$. The properties and performance of this modified Kalman filter algorithm are discussed in Monfort et al. (2017), among others.

[8] This is also done in the Affine term structure literature.

The system (10)-(11) gives the state-space representation of our affine-Q framework. The marginal moments (mean and variance) of the latent vector are used to initialize the filter, by setting $F_{0|0} = E(F_t)$ and $P_{0|0} = Var(F_t)$. Now, consider that $F_{t|t}$ and $P_{t|t}$ are available at a generic iteration t. Then, the filter proceeds recursively through the forecasting step

$$\begin{cases} F_{t+1|t} = \Phi_0 + \Phi_1 F_{t|t} \\ P_{t+1|t} = \Phi_1 P_{t|t} \Phi_1' + \Sigma\left(F_{t|t}\right) \\ CUM^Q_{t+1|t} = \Gamma_0 + \Gamma_1 F_{t+1|t} \\ M_{t+1|t} = \Gamma_1 P_{t+1|t} \Gamma_1' + \Omega, \end{cases} \tag{12}$$

and the updating step

$$\begin{cases} F_{t+1|t+1} = \left[F_{t+1|t} + P_{t+1|t}\Gamma_1' M^{-1}_{t+1|t}\left(CUM^Q_{t+1} - CUM^Q_{t+1|t}\right)\right]_+, \\ P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t}\Gamma_1' M^{-1}_{t+1|t}\Gamma_1 P_{t+1|t}, \end{cases} \tag{13}$$

where $[F]_+$ returns a vector whose i^{th} element is $\max(F_i, 0)$. This additional condition ensures that latent factor estimates remain positive for all iterations, a crucial property for stochastic volatility factors that cannot assume negative values.[9] Finally, one can construct a Gaussian quasi log-likelihood

$$-\frac{1}{2}\sum_{t=1}^{T}\left[\ln\left((2\pi)^{3J}\det\left(M_{t|t-1}\right)\right) + \left(CUM^Q_t - CUM^Q_{t|t-1}\right)' M^{-1}_{t|t-1}\left(CUM^Q_t - CUM^Q_{t|t-1}\right)\right], \tag{14}$$

and maximize it over the model parameter set.

3.2 Inference

We carry out the estimation by quasi-maximum likelihood (QML), where an approximated log-likelihood function is built within a modified version of the Kalman filter. The slight departure from a standard Kalman filter algorithm stems from the fact that the covariance matrix $\Sigma(F_t)$ depends on the latent vector of factors F_t, and is estimated by $\Sigma\left(F_{t|t}\right)$. Moreover, the distribution of the measurement error vector ϑ_t is assumed to be i.i.d. Gaussian. Note that an alternative positive distribution – such as the Gamma law – may be considered, especially for the observation equation of the second cumulant (variance). The finite sample properties of the modified Kalman

[9] Latent factors are often persistent, and therefore the algorithm tends to yield null filtered values that are clustered. This may be at odds with standard distributional assumptions that usually preclude strictly positive model-implied probabilities of staying at the same value for consecutive periods.

filter have been studied for multi-factor affine term structure interest rate models by De Jong (2000) and Kim and Singleton (2012), among others. Their Monte Carlo experiments suggest that the estimated model parameters and factors are well behaved. We refer the readers to Monfort et al. (2017) for further details on the inference.

4 Illustrative example: the Andersen et al. (2015b) model

To provide a practical illustration, we use our methodology to estimate the AFT model. This is a state-of-the-art option valuation framework that has been shown to successfully match several salient features of option panels.

4.1 The specification

In the three-factor jump-diffusive stochastic volatility model of Andersen et al. (2015b), the underlying asset price evolves according to the following general dynamics (under Q):

$$\frac{dX_t}{X_{t-}} = (r_t - \delta_t)\,dt + \sqrt{V_{1t}}dW_{1t}^Q + \sqrt{V_{2t}}dW_{2t}^Q + \eta\sqrt{V_{3t}}dW_{3t}^Q + \int_{R^2}\left(e^x - 1\right)\tilde{\mu}\left(dt, dx, dy\right), \quad (15)$$

$$dV_{1t} = \kappa_1\left(v_1 - V_{1t}\right)dt + \sigma_1\sqrt{V_{1t}}dB_{1t}^Q + \mu_1\int_{R^2} x^2 1_{\{x<0\}}\mu\left(dt, dx, dy\right), \quad (16)$$

$$dV_{2t} = \kappa_2\left(v_2 - V_{2t}\right)dt + \sigma_2\sqrt{V_{2t}}dB_{2t}^Q, \quad (17)$$

$$dV_{3t} = -\kappa_3 V_{3t}dt + \mu_3\int_{R^2}\left[\left(1-\rho_3\right)x^2 1_{\{x<0\}} + \rho_3 y^2\right]\mu\left(dt, dx, dy\right), \quad (18)$$

where $\left(W_{1t}^Q, W_{2t}^Q, W_{3t}^Q, B_{1t}^Q, B_{2t}^Q\right)$ is a five-dimensional Brownian motion with $corr\left(W_{1t}^Q, B_{1t}^Q\right) = \rho_1$ and $corr\left(W_{2t}^Q, B_{2t}^Q\right) = \rho_2$, while the remaining Brownian motions are mutually independent. The risk-neutral compensator for the jump measure μ is

$$\nu_t^Q\left(dx, dy\right) = \left\{\left(c^- 1_{\{x<0\}}\lambda_- e^{-\lambda_-|x|} + c^+ 1_{\{x>0\}}\lambda_+ e^{-\lambda_+|x|}\right)1_{\{y=0\}} + c^- 1_{\{x=0, y<0\}}\lambda_- e^{-\lambda_-|y|}\right\}dx \otimes dy,$$

$$c^- = c_0^- + c_1^- V_{1t-} + c_2^- V_{2t-} + c_3^- V_{3t-}, \quad c^+ = c_0^+ + c_1^+ V_{1t-} + c_2^+ V_{2t-} + c_3^+ V_{3t-}.$$

The AFT model clearly extends several existing one- and two-factor specifications.[10] Namely, it allows the jump tail intensity to be governed by a third factor that is distinct from – though possibly related to – market volatility. This realistic feature plays a pivotal role in explaining the observed asymmetric behavior of the priced right-versus-left jump tail risk, with the latter displaying more pronounced and persistent dynamics.

[10] The model discussed in Andersen et al. (2015a) is obtained by setting $\eta = 0$.

4.2 The conditional characteristic function

The conditional characteristic function, which provides a complete description of an asset's distribution, is very useful for the modeling and estimation of contingent claims written on that asset. It also provides a way to compute the analytical expressions of risk-neutral cumulants. Duffie et al. (2000) and Duffie et al. (2003) present a meticulous discussion on a wide range of valuation and econometric applications with conditional characteristic functions in the context of affine jump-diffusion processes.

We start from the log-price process $\{y_t\}_{t\geq 0}$ in the AFT model. The conditional characteristic function formula is

$$E_t^Q\left[e^{u(y_{t+\tau}-y_t)}\right] = \exp\left\{\alpha(u,\tau) + \beta_1(u,\tau)V_{1t} + \beta_2(u,\tau)V_{2t} + \beta_3(u,\tau)V_{3t}\right\}, \tag{19}$$

where $T = t+\tau$ is the expiration date, u is a complex number, and the functions $\alpha(u,\tau)$, $\beta_1(u,\tau)$, $\beta_2(u,\tau)$, and $\beta_3(u,\tau)$ are solutions to the following ODEs:

$$\begin{aligned}
\frac{\partial\alpha(u,\tau)}{\partial\tau} &= u\left[r-\delta-c_0^-\left(\Theta^{nc}(u,0,0)-1\right)-c_0^+\left(\Theta^{p}(u)-1\right)\right]+\beta_1\kappa_1v_1+\beta_2\kappa_2v_2 \\
&\quad + c_0^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_0^-\left(\Theta^{ni}(\beta_3)-1\right)+c_0^+\left(\Theta^{p}(u)-1\right), \qquad (20)\\
\frac{\partial\beta_1(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}-c_1^-\left(\Theta^{nc}(u,0,0)-1\right)-c_1^+\left(\Theta^{p}(u)-1\right)\right]-\beta_1\kappa_1+\frac{1}{2}u^2+\frac{1}{2}\sigma_1^2\beta_1^2+\beta_1u\sigma_1\rho_1 \\
&\quad + c_1^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_1^-\left(\Theta^{ni}(\beta_3)-1\right)+c_1^+\left(\Theta^{p}(u)-1\right), \qquad (21)\\
\frac{\partial\beta_2(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}-c_2^-\left(\Theta^{nc}(u,0,0)-1\right)-c_2^+\left(\Theta^{p}(u)-1\right)\right]-\beta_2\kappa_2+\frac{1}{2}u^2+\frac{1}{2}\sigma_2^2\beta_2^2+\beta_2u\sigma_2\rho_2 \\
&\quad + c_2^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_2^-\left(\Theta^{ni}(\beta_3)-1\right)+c_2^+\left(\Theta^{p}(u)-1\right), \qquad (22)\\
\frac{\partial\beta_3(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}\eta^2-c_3^-\left(\Theta^{nc}(u,0,0)-1\right)-c_3^+\left(\Theta^{p}(u)-1\right)\right]-\beta_3\kappa_3+\frac{1}{2}u^2\eta^2 \\
&\quad + c_3^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_3^-\left(\Theta^{ni}(\beta_3)-1\right)+c_3^+\left(\Theta^{p}(u)-1\right), \qquad (23)
\end{aligned}$$

with

$$\begin{aligned}
\Theta^{nc}(q_0,q_1,q_3) &= \int_{-\infty}^{0} e^{q_0z+q_1\mu_1z^2+q_3(1-\rho_3)\mu_3z^2}\lambda_-e^{\lambda_-z}dz, \qquad (24)\\
\Theta^{ni}(q_3) &= \int_{-\infty}^{0} e^{q_3\rho_3\mu_3z^2}\lambda_-e^{\lambda_-z}dz, \qquad (25)\\
\Theta^{p}(q_0) &= \int_{0}^{+\infty} e^{q_0z}\lambda_1e^{-\lambda_+z}dz. \qquad (26)
\end{aligned}$$

At this point, a few important comments are in order. The ODEs 20 to 23 cannot be solved analytically. Alternatively, a numerical resolution involves several challenges and sources of errors, including – but not limited to – discrete approximations of differences and high-dimensional numerical integrals. Hence, option valuation methods that rely on resolving these ODEs are computationally challenging.

Duffie et al. (2000) provide additional insights on the challenges that arise when solving these ODEs (see Equations 2.5 and 2.6 in Duffie et al. (2000) and discussion thereafter, page 1351). Simply put, in the case of pure diffusion (no jump) processes, explicit solutions to the ODEs of the conditional characteristic function can be derived. In the presence of jumps, finding closed-form solutions depends on the nature of the jump. Thus, selecting a jump distribution with an explicitly known or tractable jump transform has a practical advantage. For other process specifications, featuring, for instance, self-exciting jumps, the analytical solutions to these ODEs are either more involved or not available.

4.3 The conditional risk-neutral cumulant

In contrast to estimation methods that "directly" solve the ODEs of conditional characteristic function, implementing our risk-neutral moment approach does not require solving for $\alpha(u,\tau)$ and $\beta_i(u,\tau)$ in the conditional characteristic function, but rather entails computing $\frac{\partial^n \alpha(0,\tau)}{\partial u^n} \equiv \frac{\partial^n \alpha(u,\tau)}{\partial u^n}\Big|_{u=0}$, and $\frac{\partial^n \beta_i(0,\tau)}{\partial u^n} \equiv \frac{\partial^n \beta_i(u,\tau)}{\partial u^n}\Big|_{u=0}$, for $i = 1,2,3$, and $n = 2,3,4$. For the AFT model, we show that the partial derivatives $\frac{\partial^n \alpha(0,\tau)}{\partial u^n} \equiv \frac{\partial^n \alpha(u,\tau)}{\partial u^n}\Big|_{u=0}$, and $\frac{\partial^n \beta_i(0,\tau)}{\partial u^n} \equiv \frac{\partial^n \beta_i(u,\tau)}{\partial u^n}\Big|_{u=0}$, for $i = 1,2,3$, and $n = 2,3,4$, are explicit solutions to other ODEs given in Appendix B. The online Appendix collects the formal derivation steps to solve these ODEs analytically.

Proposition 1 *We establish that although the conditional characteristic function in the AFT model is not available in closed-form, its n^{th}-order derivative with respect to u, evaluated at 0, which gives the n^{th}-order risk-neural cumulant, has a closed-form solution.*

This highlights one of the main advantages of our risk-neutral moment-based estimation approach, especially when dealing with rich multi-factor pricing models. To better understand the result in Proposition 1, it is useful to notice that the coefficients $\alpha(u,\tau)$ and $\beta_i(u,\tau)$ associated with the conditional characteristic function have an extra argument (u), which complicates the resolution of their ODEs. Interestingly, (1) u is set to 0 when solving for the ODEs of the risk-neutral moments,

and (2) at $u = 0$, the coefficients $\alpha(u,\tau)$ and $\beta_i(u,\tau)$ in the characteristic function are also null. These properties allow us to derive a closed-form analytical solution for the ODEs of the risk-neutral moments.

To deepen our understanding of the AFT model, we illustrate the impacts of different sources of leverage on the weighting coefficients (loadings) of the factors governing the risk-neutral cumulants. In Figure 2, we observe that the leverage effect originating from the first volatility factor ($\rho_1 < 0$) increases monotonically the factor loadings of the second cumulant, yet reduces that of the third cumulant. The implication of a negative $\rho_1 < 0$ value for the fourth cumulant is broadly positive with a mix of left-skewed parabolic and linear effects. Thus, the information content of the second, third and fourth cumulants is useful to pin down the empirical value of ρ_1. Moreover, Figure 3 reveals that, while the leverage effect induced by the second volatility factor ($\rho_2 < 0$) has virtually no impact on the second and third cumulant, it produces a complementary effect on the fourth cumulant. This suggests that the fourth cumulant plays a key role in the identification of ρ_2.

4.4 Generalization of Proposition 1

Our analytical resolution of the ODEs for the risk-neutral moments is not specific to the AFT model. We now present the general expressions of the risk-neutral moments ODEs for a generic N-factor affine model, nesting the AFT specification.

Proposition 2 *Let $\beta = (\beta_j(u,\tau))'_{j=1,...N}$ denote a vector of the slope functionals associated with the characteristic function. We derive a closed-form solution for the following general risk-neutral moments ODEs:*

$$\frac{\partial\left[\frac{\partial^n \alpha}{\partial u^n}(0,\tau)\right]}{\partial \tau} = \alpha^{(n)}(\tau) + A'_\alpha \frac{\partial^n \beta}{\partial u^n}(0,\tau), \quad \frac{\partial\left[\frac{\partial^n \beta}{\partial u^n}(0,\tau)\right]}{\partial \tau} = B^{(n)}(\tau) + A\frac{\partial^n \beta}{\partial u^n}(0,\tau),$$

where $\alpha^{(n)}(\tau)$ and $B^{(n)}(\tau)$, for $n = 2,3,4$, are given in Appendix C.

The online Appendix provides a general characterization of the solution for these risk-neutral moments ODEs. Note that the steps for solving these general ODEs are exactly similar to the AFT model case.

The risk-neutral moments ODEs in Duffie et al. (2000) can be cast within the general ODEs described in Proposition 2. Hence, closed-form expressions for risk-neutral moments can be found

in the affine family described in Duffie et al. (2000). Furthermore, the results in Proposition 2 go far beyond the affine class of models discussed in Duffie et al. (2000), since the AFT model is not nested within the Duffie et al. (2000) framework.

4.5 The discretized space-state system

To estimate the AFT model, we derive its discretized state-space representation and use the modified Kalman filter described earlier. Note that the measurement equation is the same as in Equation 10, and therefore is skipped here to save space. We show in Appendix D (along with other technical details) that the transition equations for the three factors in the AFT model can be stated as

$$V_{t+1} = \Phi_0 + \Phi_1 V_t + \varepsilon_{t+1}, \tag{27}$$

where

$$\Phi_0 \equiv \Delta t \begin{pmatrix} \kappa_1 c_1 + \mu_1 \lambda_- c_0^- \\ \kappa_2 c_2 \\ \mu_3 \lambda_- c_0^- \end{pmatrix}, \quad \Phi_1 \equiv I_3 + K_1, \quad K_1 = \Delta t \begin{bmatrix} -\kappa_1 + \mu_1 \lambda_- c_1^- & \mu_1 \lambda_- c_2^- & \mu_1 \lambda_- c_3^- \\ 0 & -\kappa_2 & 0 \\ \mu_3 \lambda_- c_1^- & \mu_3 \lambda_- c_2^- & -\kappa_3 + \mu_3 \lambda_- c_3^- \end{bmatrix},$$

I_3 is a 3×3 identity matrix, $\lambda_- = 2/\lambda_-^2$, and Δt is set to $1/252$ to reflect a daily time step. In the AFT model, the latent factor F_{t+1} is a 3×1 vector denoted by $V_{t+1} \equiv (V_{1t+1}, V_{2t+1}, V_{3t+1})'$. Moreover, the transition noise is $\varepsilon_{t+1} \equiv (\varepsilon_{1t+1}, \varepsilon_{2t+1}, \varepsilon_{3t+1})'$, with a conditional covariance matrix

$$\Sigma(V_t) \equiv Var_t(\varepsilon_{t+1}) = \Delta t \begin{bmatrix} \sigma_1^2 V_{1t} + \mu_1^2 \lambda_-^* c_t^- & 0 & \mu_1 \mu_3 (1 - \rho_3) \lambda_-^* c_t^- \\ 0 & \sigma_2^2 V_{2t} & 0 \\ \mu_1 \mu_3 (1 - \rho_3) \lambda_-^* c_t^- & 0 & \mu_3^2 \left[(1 - \rho_3)^2 + \rho_3^2 \right] \lambda_-^* c_t^- \end{bmatrix}, \tag{28}$$

where $\lambda_-^* = 24/\lambda_-^4$. It is immediately apparent that $\varepsilon_{t+1} = \Sigma(V_t)^{1/2} \xi_{t+1}$, where ξ_{t+1} is the zero-mean standardized noise in Equation 11.

The initial filtering conditions are given by the unconditional moments of V_t. Namely, we set $V_{0|0} = -K_1^{-1} \Phi_0$, and $vec\left(P_{0|0}\right) = (I_9 - \Phi_1 \otimes \Phi_1)^{-1} vec\left(\Sigma\left(V_{0|0}\right)\right)$, where I_9 is a 9×9 identity matrix, and $\otimes$ is the Kronecker product. The filtering recursions and the quasi-maximum likelihood optimization are then performed to obtain factor and parameter estimates.

5 Empirical analysis

5.1 Option-implied risk-neutral moments

Our empirical analysis hinges on the construction of model-free risk-neutral cumulants for various maturities. We extract risk-neutral second, third, and fourth cumulants from a panel of European options written on the S&P500 index. The daily observations span the period September 03, 1996, through December 30, 2011, and are retrieved from OptionMetrics, a data supplier for the US option markets. Our full sample includes 3,073 trading days.[11] A detailed description of the option contracts used to construct the risk-neutral cumulants is provided in the online Appendix, which also documents a rich set of cross-maturity linear dependencies and commonalities among risk-neutral cumulants. Moreover, the online Appendix outlines a two-step principal component analysis (PCA) procedure that allows us to reduce the dimension of risk-neutral cumulants observed at different maturities.

In our empirical analysis, we employ the second, third, and fourth risk-neutral cumulants observed for $J = 8$ different maturities corresponding to 1, 2, 3, 6, 9, 12, 18, 24 months. Thus, on each observation day t, we have $3J = 24$ observations of risk-neutral moments.

To provide numbers with comparable magnitude, we report descriptive statistics for risk-neutral moments (volatility, skewness, and kurtosis) rather than second, third, and fourth cumulants. Figure 4 displays some time series properties of the risk-neutral volatility, skewness, and kurtosis for 1-month, 6-month, and 12-month maturities. A few remarks emerge when looking at the dynamics of risk-neutral moments over our sample period, which starts on September 03, 1996, and ends on December 30, 2011. First, risk-neutral volatilities are high during the 1997-2003 portion of the sample, as documented in Bollerslev et al. (2009). They are markedly higher from 2008 to 2011. The distinct spikes in volatilities (up to 70%) at the end of the sample (2008-2011) can be related to the substantial uncertainty anticipated by investors during the Great Recession. Second, the option-implied skewness is negative and decreases over the post-2008 period. This empirical evidence echoes market concerns of negative investment prospects in the wake of financial meltdowns. Drawing on a similar intuition, recent studies by Kozhan et al. (2014) and Feunou et al. (forthcoming) highlight the ability of skewness risk measures to attract premia.[12] Third, the risk-

[11] In their empirical study, AFT use Wednesday option contracts from January 01, 1996, to July 21, 2010, which yields 760 daily observations.

[12] Amaya et al. (2015) use alternative measures of realized skewness to successfully predict the cross-section of weekly returns.

neutral kurtosis takes positive values and increases over the last portion of the sample, when the likelihood of tail events becomes higher. The observed patterns remain robust through maturities, even though longer tenors tend to exhibit less fluctuation.

The descriptive values presented in Table 1 support all the empirical regularities discussed above. In Panel A of Table 1, we report for each maturity (1 to 24 months) the mean, the standard deviation, and the first-lagged autocorrelation of observed risk-neutral volatility series. The average risk-neutral volatility increases almost monotonically from 17.68% at 1-month maturity to 20.95% at 24-month maturity. As the maturity increases, option-implied volatilities become less dispersed. As expected, risk-neutral volatilities are highly persistent. First-order autocorrelations are strong and remain above 0.95 for all maturities. Detailed discussions on the memory patterns of the risk-neutral volatility can be found in Britten-Jones and Neuberger (2000) and Bollerslev et al. (2012), among others. Panel B shows the summary statistics for the third risk-neutral moment. The time series of option-implied skewness are negative on average (-1.85 to -1.62) and appear less dispersed at medium maturities (6 to 12 months). It emerges from Panel C that risk-neutral kurtosis values are positive on average, ranging between 6.15 and 8.07. Finally, risk-neutral skewness and kurtosis appear less persistent as compared to the option-implied volatility.

5.2 Model fit analysis

We explore the performance of the proposed risk-neutral moment-based estimation method as compared to the AFT estimation approach along several dimensions. First, we compare the parameter estimates and the filtered factors from both estimation strategies. We also briefly discuss the computation time efficiency of our moment-based estimation methodology vis-à-vis the AFT estimation approach for a large panel of options. Second, we assess how well the model-implied risk-neutral volatility, skewness, and kurtosis replicate their model-free counterparts. In our performance analysis, we use risk-neutral moments (volatility, skewness, and kurtosis) instead of cumulants to ease comparability with similar quantities in the literature. Third, we appraise whether the moment-based estimation results induce realistic option prices.

5.2.1 Parameter estimates, filtered factors, and implied moments

The fitted parameters from the risk-neutral moment-based and the AFT estimation approaches are reported in Table 2. Interestingly, we see that both approaches yield similar estimated parameter values for the AFT model. The parameters are also fitted with good accuracy, as evidenced by small

standard errors. Moreover, the estimated figures are economically sensible, with the expected signs and magnitudes. Note that, in our empirical investigation, some estimated parameter values may differ from those reported in AFT. Indeed, our panel includes option contracts recorded on every trading day (3,073 days from September 03, 1996, to December 30, 2011), whereas AFT's panel includes options sampled every Wednesday (760 days over January 01, 1996, to July 21, 2010).

Looking at the dynamics of the latent factors from both estimation strategies displayed in Figure 5, we notice that they evolve similarly over the observation window, spiking in the late 1990s, and markedly so during the 2009-2010 crisis period. The filtered factor series remain positive, consistent with the theoretical predictions. This observation further reflects a good performance of our moment-based estimation strategy.

As we run our estimation for all (570,108) contracts in our option panel, the computational time is an important aspect of the estimation efficiency. The risk-neutral moment-based estimation clearly dominates the AFT estimation approach in this regard. Namely, the moment-based estimation takes only 3 minutes to converge, whereas the AFT estimation runs for several days.

We now assess the ability of the moment-based estimation strategy to generate realistic model-implied risk-neutral moments. To this end, we turn to Figure 6 that shows the observed series along with the AFT estimation and the moment-based estimation implied risk-neutral volatility, skewness, and kurtosis. The observed series are constructed from the nonparametric procedure in Equation 8. The fitted parameter values in the second column of Table 2 are used to generate the risk-neutral series from the moment-based estimation. The corresponding series from the AFT estimation are constructed using the values in the fourth column of Table 2. Qualitatively, we see that the moment-based estimation approach implies risk-neutral moment paths that track well those from the AFT estimation strategy as well as their nonparametric counterparts.

In a nutshell, the aforementioned empirical observations suggest that the risk-neutral moment-based estimation approach is rather fast, is easy to implement, and delivers accurate parameter and factor estimates.

5.2.2 Matching the option prices

We now gauge the ability of the risk-neutral moment-based estimation approach to yield empirically grounded prices. We compute the ratio of AFT estimation to risk-neutral moment-based estimation VWRMSEs. A ratio exceeding (resp. below) 1 will indicate that the risk-neutral

moment-based estimation approach yields a smaller (resp. larger) error on average, and thus, tends to outperform (resp. underperform) the AFT estimation in matching the observed option values.

Table 3 gives the pricing errors, "disaggregated" by moneyness and maturity. Overall, the model fit generates a VWRMSE of 3.65% for the risk-neutral moment-based estimation, a pricing error which is fairly close to that of the AFT estimation (3.09%). Across the board, the VWRMSE ratio between the two estimation approaches is above 0.75, and often close to 1. Thus, the risk-neutral moment-based estimation strategy delivers realistic option values as compared to the AFT estimation.

5.3 Risk premia

In this section, we assess the risk premia implied by our risk-neutral moment-based estimation approach. To this end, we now cast the factor dynamics under the physical probability measure. From an implementation perspective, this is done by allowing for "free" or shifted physical drift parameters κ_1^P, v_1^P, κ_2^P, and v_2^P in the expressions of Φ_0 and Φ_1 in Equation 27. The physical parameters, which are distinct from their risk-neutral counterparts κ_1^Q, v_1^Q, κ_2^Q, and v_2^Q appearing in the measurement Equation 10 of risk-neutral cumulants, are additional parameters that can be easily estimated in our framework. These two sets of parameters induce wedges between risk-neutral and physical quantities, thus reflecting risk premia. As pointed out by Andersen et al. (2015b), common structure-preserving transformations restrict other model parameters such as σ_1, σ_2, η, ρ_1, ρ_2, ρ_3, μ_1, and κ_3 to be identical under both physical and risk-neutral probability measures.[13]

Table 4 shows that the drift parameters are estimated with accuracy, as the fitted values are all significant at 10% level. Moreover, Figure 7 displays the paths of implied risk-neutral (solid lines) and physical (dashed lines) moments. The shaded areas between the risk-neutral and physical series reflect the corresponding premia. We observe that the implied variance risk premium (defined as the risk-neutral minus the physical series) is mostly positive, consistent with the empirical evidence documented, for instance, in Bollerslev et al. (2009). By contrast, the model-implied skewness risk premium is mostly negative and in line with the results in Feunou et al. (forthcoming) and Kozhan et al. (2014), among others. Finally, the evidence for the model-implied kurtosis risk premium is mixed, probably due to sampling variation.

[13] To ensure tractability, it is often assumed that the model class is identical and affine under both measures, a feature referred to as structure-preserving transformations.

6 Conclusion

We propose a new method to identify the latent factors and estimate parameters of affine continuous time stochastic volatility models. This procedure "collapses" raw option prices into risk-neutral cumulants, and then uses these conditional cumulants – instead of option prices – in the fitting step. The adopted approach can be viewed as a portfolio-based estimation strategy where the computation of risk-neutral cumulants amounts to constructing portfolios of European option contracts, weighted across moneyness.

The proposed methodology is well capable of handling a large data set of options without any discretionary selection of observations. It also helps circumvent several sources of errors rooted in the numerical approximations of the characteristic function and its transform inversion. Moreover, the affine mapping between each cumulant and the latent factors allows for a straightforward implementation of a linear Kalman filtering algorithm. Using the AFT model, we show that our risk-neutral moment-based estimation approach generates parameter estimates, filtered factors, risk-neutral moments, and option pricing errors that are similar to those obtained from the AFT estimation method. Interestingly, the proposed risk-neutral moment-based estimation strategy is effective, easy to implement, and robust, as compared to the AFT estimation approach.

References

Aït-Sahalia, Yacine, Karaman, Mustafa, and Mancini, Loriano (2015), "The term structure of variance swaps and risk premia," *Working Paper, Princeton University, NBER, Swiss Finance Institute*.

Amaya, Diego, Christoffersen, Peter, Jacobs, Kris, and Vasquez, Aurelio (2015), "Does realized skewness predict the cross-section of equity returns?" *Journal of Financial Economics*, 118(1), 135–167.

Amengual, Dante, and Xiu, Dacheng (2015), "Resolution of policy uncertainty and sudden declines in volatility," *Working Paper, CEMFI and Chicago Booth*.

Andersen, Torben, Fusari, Nicola, and Todorov, Viktor (2015a), "Parametric inference and dynamic state recovery from option panels," *Econometrica*, 83(3), 1081–1145.

_ (2015b), "The risk premia embedded in index options," *Journal of Financial Economics*, 117(3), 558–584.

Andersen, Torben G., Benzoni, Luca, and Lund, Jesper (2002), "An empirical investigation of continuous-time equity return models," *Journal of Finance*, 57, 1239–1284.

Bakshi, Gurdip, and Madan, Dilip (2000), "Spanning and derivative security valuation," *Journal of Financial Economics*, 55(2), 205–238.

Bakshi, Gurdip, Cao, Charles, and Chen, Zhiwu (1997), "Empirical performance of alternative option pricing models," *Journal of Finance*, 52(5), 2003–2049.

Bakshi, Gurdip, Kapadia, N., and Madan, Dilip (2003), "Stock return characteristics, skew laws and the differential pricing of individual equity options," *Review of Financial Studies*, 16(1), 101–143.

Bates, David (2000), "Post-'87 crash fears in the S&P 500 futures option market," *Journal of Econometrics*, 96(1–2), 181–238.

_ (2006), "Maximum likelihood estimation of latent affine processes," *Review of Financial Studies*, 26(9), 909–965.

Bollerslev, Tim, Sizova, Natalia, and Tauchen, George (2012), "Volatility in equilibrium: Asymmetries and dynamic dependencies," *Review of Finance*, 16(1), 31–80.

Bollerslev, Tim, Tauchen, George, and Zhou, Hao (2009), "Expected stock returns and variance risk premia," *Review of Financial Studies*, 22(11), 4463–4492.

Britten-Jones, Mark, and Neuberger, Anthony (2000), "Option prices, implied price processes, and stochastic volatility," *Journal of Finance*, 55(2), 839–866.

Carr, Peter, and Wu, Liuren (2007), "Stochastic skew in currency options," *Journal of Financial Economics*, 86(1), 213–247.

Carr, Peter, Lee, Roger, and Wu, Liuren (2012), "Variance swaps on time-changed lévy processes," *Finance and Stochastics*, 16, 335–355.

Chang, Bo, Christoffersen, Peter, Jacobs, Kris, and Vainberg, Gregory (2012), "Option-implied measures of equity risk," *Review of Finance*, 16(2), 385–428.

Chernov, Mikhail, and Ghysels, Eric (2000), "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation," *Journal of Econometrics*, 56, 407–458.

Christoffersen, Peter, Heston, Steven, and Jacobs, Kris (2009), "The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well," *Management Science*, 55(12), 1914–1932.

Darolles, Serge, Gourieroux, Christian, and Jasiak, Joann (2006), "Structural Laplace transform and compound autoregressive models," *Journal of Time Series Analysis*, 27(4), 477–503.

De Jong, Frank (2000), "Time series and cross-section information in affine term-structure models," *Journal of Business and Economic Statistics*, 18(3), 300–314.

Duan, Jin-Chuan, and Yeh, Chung-Ying (2010), "Jump and volatility risk premiums implied by VIX," *Journal of Economic Dynamics and Control*, 34, 2232–2244.

Duffie, Darrell, Filipovic, Damir, and Schachermayer, Walter (2003), "Affine processes and applications in finance," *Annals of Applied Probability*, 13(3), 984–1053.

Duffie, Darrell, Pan, Jun, and Singleton, Kenneth (2000), "Transform analysis and option pricing for affine jump-diffusions," *Econometrica*, 68, 1343–1377.

Egloff, Daniel, Leippold, Markus, and Wu, Liuren (2010), "The term structure of variance swap rates and optimal variance swap investments," *Journal of Financial and Quantitative Analysis*, 45(5), 1279–1310.

Eraker, Bjørn (2004), "Do stock prices and volatility jump? Reconciling evidence from spot and option prices" *Journal of Finance*, 59(3), 1367–1404.

Feunou, Bruno, and Tédongap, Roméo (2012), "A stochastic volatility model with conditional skewness," *Journal of Business and Economic Statistics*, 30(4), 576–591.

Feunou, Bruno, Jahan-Parvar, Mohammad R., and Okou, Cédric (forthcoming), "Downside variance risk premium," *Journal of Financial Econometrics*.

Gagliardini, Patrick, Gourieroux, Christian, and Renault, Eric (2011), "Efficient derivative pricing by the extended method of moments," *Econometrica*, 79, 1181–1232.

Gallant, Ronald, and Tauchen, George (1996), "Which moments to match?" *Econometric Theory*, 12(4), 657–681.

Heston, Steven (1993), "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *Review of Financial Studies*, 6(2), 327–343.

Huang, Jing-Zhi, and Wu, Liuren (2004), "Specification analysis of option pricing models based on time-changed lévy processes," *Journal of Finance*, 59(3), 1405–1439.

Johannes, Michael, Polson, Nicholas, and Stroud, Jonathan (2009), "Optimal filtering of jump-diffusions: Extracting latent states from asset prices," *Review of Financial Studies*, 22, 2759–2799.

Jones, Christopher (2003), "The dynamics of stochastic volatility," *Journal of Econometrics*, 116, 181–224.

Joslin, Scott, Singleton, Kenneth, and Zhu, Haoxiang (2011), "A new perspective on Gaussian dynamic term structure models," *Review of Financial Studies*, 24(3), 926–970.

Kaeck, Andreas, and Alexander, Carol (2012), "Volatility dynamics for the S&P 500: Further evidence from non-affine, multi-factor jump diffusions," *Journal of Banking and Finance*, 36(11), 3110–3121.

Kim, Don H., and Singleton, Kenneth J. (2012), "Term structure models and the zero bound: An empirical investigation of Japanese yields," *Journal of Econometrics*, 170(1), 32–49.

Kozhan, Roman, Neuberger, Anthony, and Schneider, Paul (2014), "The skew risk premium in the equity index market," *Review of Financial Studies*, 26(9), 2174–2203.

Monfort, Alain, Pegoraro, Fulvio, Renne, Jean-Paul, and Roussellet, Guillaume (2017), "Staying at zero with affine processes: An application to term structure modelling," *Journal of Econometrics*, 201(2), 348 – 366.

Pan, Jun (2002), "The jump-risk premia implicit in options: Evidence from an integrated time-series study," *Journal of Financial Economics*, 63, 3–50.

Renault, Eric, (1997) *"Econometric Models of Option Pricing Errors," in Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, eds. D.M. Kreps and K.F. Wallis, Cambridge: Cambridge University Press, pp. 223-278.

Singleton, Kenneth (2006) *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*, Princeton: Princeton University Press.

Van der Merwe, Rudolph, and Wan, Eric (2001), "The square-root unscented Kalman filter for state and parameter-estimation," *International Conference on Acoustics, Speech and Signal Processing*, pp. 3461–3464.

Appendices

Appendix A: Expressions of location and slope coefficients of cumulants for the Heston (1993) model

We present the expressions of the location $A_\tau^{(n)}$ and slope $B_\tau^{(n)}$ coefficients in the factor representation of the n^{th}-order cumulant ($CUM_\tau^{(n)}, n = 2, 3, 4$) in Equation 9 at a given maturity $\tau = T - t$ for the Heston (1993) model.

Second Cumulant ($n = 2$) We have

$$A_\tau^{(2)} = \frac{\left(\begin{array}{c} -8b^2\rho\sigma\tau\,\mathrm{e}^{b\tau} - 8b^2\rho\sigma\tau\,\mathrm{e}^{2b\tau} + 8b^3\tau\,\mathrm{e}^{2b\tau} + 4b\sigma^2\tau\,\mathrm{e}^{b\tau} + 2b\sigma^2\tau\,\mathrm{e}^{2b\tau} \\ -16b\rho\sigma\,\mathrm{e}^{b\tau} + 16b\rho\sigma\,\mathrm{e}^{2b\tau} + 8b^2\mathrm{e}^{b\tau} - 8b^2\mathrm{e}^{2b\tau} + 4\sigma^2\mathrm{e}^{b\tau} - 5\sigma^2\mathrm{e}^{2b\tau} + \sigma^2 \end{array} \right) a\mathrm{e}^{-2b\tau}}{8b^4},$$

and

$$B_\tau^{(2)} = \frac{\left(4b^2\rho\sigma\tau\,\mathrm{e}^{b\tau} - 2b\sigma^2\tau\,\mathrm{e}^{b\tau} + 4b\rho\sigma\,\mathrm{e}^{b\tau} - 4b\rho\sigma\,\mathrm{e}^{2b\tau} - 4b^2\mathrm{e}^{b\tau} + 4b^2\mathrm{e}^{2b\tau} + \sigma^2\mathrm{e}^{2b\tau} - \sigma^2\right)\mathrm{e}^{-2b\tau}}{4b^3}.$$

Third Cumulant ($n = 3$) We have

$$A_\tau^{(3)} = -3/2\,\frac{a\sigma\,\mathrm{e}^{-3b\tau}}{b^5}\left(\begin{array}{c} -1/2\left(-1/12\sigma^2 + \left((b\tau+2)b\rho\sigma - b^2 + \left(-1/2b\tau - 1/2\right)\sigma^2\right)\mathrm{e}^{b\tau}\right)\sigma \\ + \left(\begin{array}{c} -2\left(b\tau+2\right)b^3\rho - \left(b\tau+2\right)^2 b\rho\sigma^2 + \\ \left(b^2\rho^2\tau^2 + 6\rho^2 + \left(4\rho^2\tau + 2\tau\right)b + 2\right)b^2\sigma + \left(1/4\,b^2\tau^2 + 3/4\,b\tau + 5/8\right)\sigma^3 \end{array} \right)\mathrm{e}^{2b\tau} \\ + \left(\begin{array}{c} -2\left(b\tau-2\right)b^3\rho + 2\left(\left(\rho^2+1/2\right)b\tau - 3\rho^2 - 5/4\right)b^2\sigma \\ + \left(1/4\,b\tau - \frac{11}{12}\right)\sigma^3 + \left(-3/2\,b^2\rho\tau + 5b\rho\right)\sigma^2 \end{array} \right)\mathrm{e}^{3b\tau} \end{array} \right),$$

and

$$B_\tau^{(3)} = 3/2\frac{\sigma\,\mathrm{e}^{-3b\tau}}{b^5}\left(\begin{array}{c} -\left(-1/8\sigma^2 + \left(\left(b\tau+3/2\right)b\rho\sigma - b^2 + \left(-1/2\,b\tau - 1/4\right)\sigma^2\right)\mathrm{e}^{b\tau}\right)\sigma \\ + \left(\begin{array}{c} -\left(b\tau+2\right)b^2\rho\sigma^2\tau - 2\left(b\tau+1\right)b^3\rho \\ + \left(b^2\rho^2\tau^2 + 2\rho^2 + \left(2\rho^2\tau + 2\tau\right)b\right)b^2\sigma \\ + \left(1/4\,b^2\tau^2 + 1/4\,b\tau - 1/8\right)\sigma^3 \end{array} \right)\mathrm{e}^{2b\tau} \\ + 2\left(b\rho - \sigma/2\right)\left(-b\rho\sigma + b^2 + 1/4\,\sigma^2\right)\mathrm{e}^{3b\tau} \end{array} \right).$$

Fourth Cumulant ($n = 4$) We have

$$A_\tau^{(4)} = 12\,\frac{a\sigma^2\mathrm{e}^{-4b\tau}}{b^8}\left(\begin{array}{c} -1/16\left(-1/16\sigma^2 + \left(\left(\rho\sigma\tau - 1\right)b^2 - 1/2\sigma^2 + \left(-1/2\sigma^2\tau + 2\rho\sigma\right)b\right)\mathrm{e}^{b\tau}\right)\sigma^2 \\ + \left(\begin{array}{c} \left(1/4\,\rho^2\sigma^2\tau^2 - 1/2\,\rho\sigma\tau + 1/8\right)b^4 + \left(-1/4\,\rho\sigma^2\tau^2 + \left(\rho^2 + 3/8\right)\sigma\tau - \rho\right)b^3\sigma \\ + 5/4\,\left(1/20\sigma^2\tau^2 - \frac{7\rho\sigma\tau}{10} + \rho^2 + 3/10\right)b^2\sigma^2 - 3/4\,\left(-\frac{5\sigma\tau}{24} + \rho\right)b\sigma^3 + \frac{7\sigma^4}{64} \end{array} \right)\mathrm{e}^{2b\tau} \\ + \left(\begin{array}{c} -1/6\left(\rho\sigma\tau - 3\right)b^5\rho^2\tau^2 - \left(-1/4\,\rho^2\sigma^2\tau^2 + \left(\rho^2+1\right)\rho\sigma\tau - 2\rho^2 - 1/2\right)b^5\tau \\ + \left(-1/8\,\rho\sigma^3\tau^3 + 7/4\,\left(\rho^2 + 3/14\right)\sigma^2\tau^2 + 3\rho^2 + \left(-3\rho^3\tau - 4\rho\tau\right)\sigma + 1/2\right)b^4 \\ -4\left(-\frac{\sigma^3\tau^3}{192} + \frac{7\rho\sigma^2\tau^2}{32} + \rho^3 + \rho + \left(-5/4\,\rho^2\tau - \frac{9\tau}{32}\right)\sigma\right)b^3\sigma \\ + 5\,\left(1/40\,\sigma^2\tau^2 - \frac{7\rho\sigma\tau}{16} + \rho^2 + 3/16\right)b^2\sigma^2 - \frac{15\,b\sigma^3}{8}\left(-\frac{3\sigma\tau}{20} + \rho\right) + \frac{7\sigma^4}{32} \end{array} \right)\mathrm{e}^{3b\tau} \\ + \left(\begin{array}{c} \left(\rho^2+1/4\right)b^5\tau + \left(-\left(\rho^2+3/2\right)\rho\sigma\tau - 3\rho^2 - 5/8\right)b^4 \\ +4\left(\rho^3 + 5/4\,\rho + \left(3/8\,\rho^2\tau + \frac{3\tau}{32}\right)\sigma\right)b^3\sigma \\ -\frac{25\,b^2\sigma^2}{4}\left(1/10\,\rho\sigma\tau + \rho^2 + \frac{11}{50}\right) \\ -\frac{93\sigma^4}{256} + \left(\frac{5\sigma^4\tau}{64} + 11/4\,\rho\sigma^3\right)b \end{array} \right)\mathrm{e}^{4b\tau} \end{array} \right),$$

and

$$B_{\tau}^{(4)} = 2\,\frac{\sigma^2 \mathrm{e}^{-4\,b\tau}}{b^7}\left(\begin{array}{c} \frac{\left(-3/4\,\sigma^2+9\left(\left(\rho\,\sigma\,\tau-1\right)b^2-1/3\,\sigma^2+\left(-1/2\,\sigma^2\tau+5/3\,\rho\,\sigma\right)b\right)\mathrm{e}^{b\tau}\right)\sigma^2}{8} \\ +\left(\begin{array}{c} -3/8\,\sigma^4+\left(-3\,\rho^2\sigma^2\tau^2+6\,\rho\,\sigma\,\tau-3/2\right)b^4 \\ -9\,\left(-1/3\,\rho\,\sigma^2\tau^2+\left(\rho^2+1/2\right)\sigma\,\tau-\rho\right)b^3\sigma \\ -9\,\left(1/12\,\sigma^2\tau^2-5/6\,\rho\,\sigma\,\tau+\rho^2+1/4\right)b^2\sigma^2 \\ +\left(-\frac{9\,\sigma^4\tau}{8}+\frac{15\,\rho\,\sigma^3}{4}\right)b \end{array}\right)\mathrm{e}^{2\,b\tau} \\ +\left(\begin{array}{c} 3/8\,\sigma^4+\left(\rho^3\sigma\,\tau^3-3\,\rho^2\tau^2\right)b^6 \\ +3\,\left(-1/2\,\rho^2\sigma^2\tau^2+\left(\rho^2+2\right)\rho\,\sigma\,\tau-2\,\rho^2-1\right)b^5\tau \\ +\left(3/4\,\rho\,\sigma^3\tau^3-6\,\left(\rho^2+3/8\right)\sigma^2\tau^2+6\,\left(\rho^2+2\right)\rho\,\sigma\,\tau-6\,\rho^2\right)b^4 \\ +\left(-1/8\,\sigma^4\tau^3+3\,\rho\,\sigma^3\tau^2+6\,\rho^3\sigma+\left(-9\,\rho^2\tau-9/4\,\tau\right)\sigma^2\right)b^3 \\ +\left(-3/8\,\sigma^4\tau^2+\frac{21\,\rho\,\sigma^3\tau}{8}+\frac{9\,\sigma^2}{8}\right)b^2-\frac{\left(3/2\,\sigma\,\tau+15\,\rho\right)b\sigma^3}{8} \\ +6\,\left(-5/4\,b\rho\sigma+\left(\rho^2+1/4\right)b^2+\frac{5\,\sigma^2}{16}\right)\left(-b\rho\sigma+b^2+1/4\,\sigma^2\right)\mathrm{e}^{4\,b\tau} \end{array}\right)\mathrm{e}^{3\,b\tau} \end{array}\right).$$

Appendix B: ODEs for cumulants in the AFT model

As stated in Equation 9, the n^{th}-order conditional cumulant corresponds to the n^{th}-order partial derivative with respect to $u \in \mathbb{C}$, evaluated at $u = 0$, of the conditional log characteristic function. In the Andersen et al. (2015b) model, the conditional log characteristic function writes

$$\ln E_t^{Q}\left[e^{u(y_{t+\tau}-y_t)}\right] = \alpha(u,\tau) + \beta_1(u,\tau)V_{1t} + \beta_2(u,\tau)V_{2t} + \beta_3(u,\tau)V_{3t},$$

where the functions $\alpha(u,\tau)$, $\beta_1(u,\tau)$, $\beta_2(u,\tau)$, and $\beta_3(u,\tau)$ are the solutions to the following ODEs:

$$\begin{aligned}
\frac{\partial\alpha(u,\tau)}{\partial\tau} &= u\left[r-\delta-c_0^-\left(\Theta^{nc}(u,0,0)-1\right)-c_0^+\left(\Theta^{p}(u)-1\right)\right]+\beta_1\kappa_1\varepsilon_1+\beta_2\kappa_2\varepsilon_2\\
&\quad+c_0^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_0^-\left(\Theta^{ni}(\beta_3)-1\right)+c_0^+\left(\Theta^{p}(u)-1\right),\\
\frac{\partial\beta_1(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}-c_1^-\left(\Theta^{nc}(u,0,0)-1\right)-c_1^+\left(\Theta^{p}(u)-1\right)\right]-\beta_1\kappa_1+\frac{1}{2}u^2+\frac{1}{2}\sigma_1^2\beta_1^2+\beta_1u\sigma_1\rho_1\\
&\quad+c_1^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_1^-\left(\Theta^{ni}(\beta_3)-1\right)+c_1^+\left(\Theta^{p}(u)-1\right),\\
\frac{\partial\beta_2(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}-c_2^-\left(\Theta^{nc}(u,0,0)-1\right)-c_2^+\left(\Theta^{p}(u)-1\right)\right]-\beta_2\kappa_2+\frac{1}{2}u^2+\frac{1}{2}\sigma_2^2\beta_2^2+\beta_2u\sigma_2\rho_2\\
&\quad+c_2^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_2^-\left(\Theta^{ni}(\beta_3)-1\right)+c_2^+\left(\Theta^{p}(u)-1\right),\\
\frac{\partial\beta_3(u,\tau)}{\partial\tau} &= u\left[-\frac{1}{2}\eta^2-c_3^-\left(\Theta^{nc}(u,0,0)-1\right)-c_3^+\left(\Theta^{p}(u)-1\right)\right]-\beta_3\kappa_3+\frac{1}{2}u^2\eta^2\\
&\quad+c_3^-\left(\Theta^{nc}(u,\beta_1,\beta_3)-1\right)+c_3^-\left(\Theta^{ni}(\beta_3)-1\right)+c_3^+\left(\Theta^{p}(u)-1\right),
\end{aligned}$$

with

$$\begin{aligned}
\Theta^{nc}(q_0,q_1,q_3) &= \int_{-\infty}^{0} e^{q_0z+q_1\mu_1z^2+q_3(1-\rho_3)\mu_3z^2}\lambda_-e^{\lambda_-z}dz,\\
\Theta^{ni}(q_3) &= \int_{-\infty}^{0} e^{q_3\rho_3\mu_3z^2}\lambda_-e^{\lambda_-z}dz,\\
\Theta^{p}(q_0) &= \int_{0}^{+\infty} e^{q_0z}\lambda_+e^{-\lambda_+z}dz.
\end{aligned}$$

Therefore, we need to compute the partial derivatives with respect to $u \in \mathbb{C}$ of the functions $\alpha(u,\tau)$, $\beta_1(u,\tau)$, $\beta_2(u,\tau)$, and $\beta_3(u,\tau)$. Note that the computation of these partial derivatives is not straightforward and requires solving a system of ODEs. We stack the slope functionals in a single vector $\beta = (\beta_1,\beta_3,\beta_2)'$ to allow for a compact treatment.

Appendix B-1: First cumulant

We show that the factor loading vector for the first cumulant is the solution of the following ODE:

$$\frac{\partial\left[\frac{\partial\beta}{\partial u}(0,\tau)\right]}{\partial\tau} = B + A\frac{\partial\beta}{\partial u}(0,\tau), \tag{29}$$

with

$$B = \begin{pmatrix} -\frac{1}{2} + c_1^- \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) + c_1^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}) \\ -\frac{1}{2}\eta^2 + c_3^- \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) + c_3^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}) \\ -\frac{1}{2} + c_2^- \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) + c_2^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}) \end{pmatrix},$$

$$A = \begin{bmatrix} -\kappa_1 + c_1^- \frac{\partial \Theta^{nc}}{\partial q_1}(\mathbf{0}) & c_1^- \frac{\partial \Theta^{nc}}{\partial q_3}(\mathbf{0}) + c_1^- \frac{\partial \Theta^{nt}}{\partial q_3}(\mathbf{0}) & 0 \\ c_3^- \frac{\partial \Theta^{nc}}{\partial q_1}(\mathbf{0}) & -\kappa_3 + c_3^- \frac{\partial \Theta^{nc}}{\partial q_3}(\mathbf{0}) + c_3^- \frac{\partial \Theta^{nt}}{\partial q_3}(\mathbf{0}) & 0 \\ c_2^- \frac{\partial \Theta^{nc}}{\partial q_1}(\mathbf{0}) & c_2^- \frac{\partial \Theta^{nc}}{\partial q_3}(\mathbf{0}) + c_2^- \frac{\partial \Theta^{nt}}{\partial q_3}(\mathbf{0}) & -\kappa_2 \end{bmatrix}.$$

Hence,

$$\frac{\partial \beta}{\partial u}(0,\tau) = e^{\tau A}\left(\int_0^\tau e^{-sA} ds\right) B = \left(e^{\tau A} - I\right) A^{-1} B. \tag{30}$$

Appendix B-2: Second cumulant

We prove that the factor loading vector for the second cumulant is the solution of the following ODE:

$$\frac{\partial \left[\frac{\partial^2 \beta}{\partial u^2}(0,\tau)\right]}{\partial \tau} = B^{(2)}(\tau) + A \frac{\partial^2 \beta}{\partial u^2}(0,\tau), \tag{31}$$

with

$$B^{(2)}(\tau) = \left(B_1^{(2)}(\tau), B_3^{(2)}(\tau), B_2^{(2)}(\tau)\right)',$$

$$B_j^{(2)}(\tau) = B_{j,c}^{(2)} + B_{j,s}^{(2)\prime} \frac{\partial \beta}{\partial u}(0,\tau) + \frac{\partial \beta}{\partial u}(0,\tau)' B_{j,q}^{(2)} \frac{\partial \beta}{\partial u}(0,\tau), \quad for \quad j = 1,2,3.$$

Specifically,

$$\begin{cases} B_{1,c}^{(2)} = 1 + c_1^- \frac{\partial^2 \Theta^{nc}}{\partial q_0^2}(\mathbf{0}) + c_1^+ \frac{\partial^2 \Theta^p}{\partial q_0^2}(\mathbf{0}) - 2c_1^- \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) - 2c_1^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}) \\ B_{3,c}^{(2)} = c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_0^2}(\mathbf{0}) + c_3^+ \frac{\partial^2 \Theta^p}{\partial q_0^2}(\mathbf{0}) - 2c_3 \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) - 2c_3^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}) \\ B_{2,c}^{(2)} = 1 + c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_0^2}(\mathbf{0}) + c_2^+ \frac{\partial^2 \Theta^p}{\partial q_0^2}(\mathbf{0}) - 2c_2 \frac{\partial \Theta^{nc}}{\partial q_0}(\mathbf{0}) - 2c_2^+ \frac{\partial \Theta^p}{\partial q_0}(\mathbf{0}), \end{cases}$$

$$\begin{aligned} B_{1,s}^{(2)} &= \begin{bmatrix} 2\sigma_1\rho_1 + 2c_1^- \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) & 2c_1^- \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) & 0 \end{bmatrix}', \\ B_{3,s}^{(2)} &= \begin{bmatrix} 2c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) & 2c_3^- \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) & 0 \end{bmatrix}', \\ B_{2,s}^{(2)} &= \begin{bmatrix} 2c_2^- \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) & 2c_2^- \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) & 2\sigma_2\rho_2 \end{bmatrix}', \end{aligned}$$

$$B_{1,q}^{(2)} = \begin{bmatrix} \sigma_1^2 + c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & c_1^- \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 0 \\ c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & c_1^- \frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + c_1 \frac{\partial^2 \Theta^{nt}}{\partial q_3^2}(\mathbf{0}) & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_{3,q}^{(2)} = \begin{bmatrix} c_3^- \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 0 \\ c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & c_3^- \frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + c_3^- \frac{\partial^2 \Theta^{nt}}{\partial q_3^2}(\mathbf{0}) & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_{2,q}^{(2)} = \begin{bmatrix} c_2^- \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 0 \\ c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & c_2^- \frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + c_2 \frac{\partial^2 \Theta^{nt}}{\partial q_3^2}(\mathbf{0}) & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix}.$$

Denote

$$C_0 = A^{-1}B, \quad \varphi_{jc} = B_{j,c}^{(2)} - B_{j,s}^{(2)\prime}C_0 + C_0'B_{j,q}^{(2)}C_0, \quad \varphi_{js} = B_{j,s}^{(2)} - 2B_{j,q}^{(2)}C_0.$$

We can then write

$$B_j^{(2)}(\tau) = \varphi_{jc} + \varphi_{js}'e^{\tau A}C_0 + C_0'e^{\tau A'}B_{j,q}^{(2)}e^{\tau A}C_0, \quad for \;\; j = 1,2,3.$$

In addition, we prove that

$$\frac{\partial\left[\frac{\partial^2\alpha}{\partial u^2}(\mathbf{0},\tau)\right]}{\partial\tau} = \alpha^{(2)}(\tau) + A_\alpha'\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau), \tag{32}$$

with

$$\begin{aligned}
\alpha^{(2)}(\tau) &= \alpha_c^{(2)} + \alpha_s^{(2)\prime}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \frac{\partial\beta}{\partial u}(\mathbf{0},\tau)'\alpha_q^{(2)}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau), \\
&= \alpha_c^{(2)} - \alpha_s^{(2)\prime}C_0 + C_0'\alpha_q^{(2)}C_0 + \left(\alpha_s^{(2)} - 2\alpha_q^{(2)}C_0\right)'e^{\tau A}C_0 + C_0'e^{\tau A'}\alpha_q^{(2)}e^{\tau A}C_0, \\
&\equiv \varphi_{0c} + \varphi_{0s}'e^{\tau A}C_0 + C_0'e^{\tau A'}\alpha_q^{(2)}e^{\tau A}C_0,
\end{aligned}$$

and

$$\begin{aligned}
A_\alpha &= \left[\begin{array}{ccc} \kappa_1 c_1 + c_0\frac{\partial\Theta^{nc}}{\partial q_1}(\mathbf{0}) & c_0^-\left(\frac{\partial\Theta^{nc}}{\partial q_3}(\mathbf{0}) + \frac{\partial\Theta^{ni}}{\partial q_3}(\mathbf{0})\right) & \kappa_2 c_2 \end{array}\right]', \\
\alpha_c^{(2)} &= c_0\frac{\partial^2\Theta^{nc}}{\partial q_0^2}(\mathbf{0}) + c_0^+\frac{\partial^2\Theta^p}{\partial q_0^2}(\mathbf{0}) - 2c_0\frac{\partial\Theta^{nc}}{\partial q_0}(\mathbf{0}) - 2c_0^+\frac{\partial\Theta^p}{\partial q_0}(\mathbf{0}), \\
\alpha_s^{(2)} &= \left[\begin{array}{ccc} 2c_0\frac{\partial^2\Theta^{nc}}{\partial q_0\partial q_1}(\mathbf{0}) & 2c_0\frac{\partial^2\Theta^{nc}}{\partial q_0\partial q_3}(\mathbf{0}) & 0 \end{array}\right]', \\
\alpha_q^{(2)} &= \left[\begin{array}{ccc} c_0\frac{\partial^2\Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & c_0^-\frac{\partial^2\Theta^{nc}}{\partial q_1\partial q_3}(\mathbf{0}) & 0 \\ c_0^-\frac{\partial^2\Theta^{nc}}{\partial q_1\partial q_3}(\mathbf{0}) & c_0\frac{\partial^2\Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + c_0\frac{\partial^2\Theta^{ni}}{\partial q_3^2}(\mathbf{0}) & 0 \\ 0 & 0 & 0 \end{array}\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{\partial^2\alpha}{\partial u^2}(\mathbf{0},\tau) &= \int_0^\tau\left(\alpha^{(2)}(s) + A_\alpha'\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},s)\right)ds, \\
&= \int_0^\tau\alpha^{(2)}(s)\,ds + A_\alpha'\int_0^\tau\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},s)\,ds.
\end{aligned} \tag{33}$$

Appendix B-3: Third cumulant

Using similar strategy, we demonstrate that the factor loading vector for the third cumulant is the solution of the following ODE:

$$\frac{\partial\left[\frac{\partial^3\beta}{\partial u^3}(\mathbf{0},\tau)\right]}{\partial\tau} = B^{(3)}(\tau) + A\frac{\partial^3\beta}{\partial u^3}(\mathbf{0},\tau) \tag{34}$$

with

$$B^{(3)}(\tau) = \left(B_1^{(3)}(\tau), B_3^{(3)}(\tau), B_2^{(3)}(\tau)\right)',$$

$$
\begin{aligned}
B_{j}^{(3)}(\tau) &= B_{j,c}^{(3)} + B_{j,s}^{(3)\prime}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \frac{\partial\beta}{\partial u}(\mathbf{0},\tau)' B_{j,q}^{(3)}(\tau)\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + B_{j,h}^{(3)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau),\\
B_{j,q}^{(3)}(\tau) &= B_{jqc} + B_{jqs}(\tau), \quad B_{jqs}(\tau) = \begin{bmatrix} B_{jqs}^{(1)\prime}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & B_{jqs}^{(3)\prime}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & B_{jqs}^{(2)\prime}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau)\end{bmatrix}\\
B_{j,h}^{(3)}(\tau) &= B_{jhc} + B_{jhs}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau), \quad for \quad j = 1,2,3.
\end{aligned}
$$

Moreover,

$$
B_{jqs}^{(2)\prime} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix},
$$

$$
\begin{cases}
B_{1,c}^{(3)} = c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^3}(\mathbf{0}) + c_1^{+}\frac{\partial^3\Theta^{p}}{\partial q_0^3}(\mathbf{0}) - 3c_1\frac{\partial^2\Theta^{nc}}{\partial q_0^2}(\mathbf{0}) - 3c_1^{+}\frac{\partial^2\Theta^{p}}{\partial q_0^2}(\mathbf{0})\\
B_{3,c}^{(3)} = c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^3}(\mathbf{0}) + c_3^{+}\frac{\partial^3\Theta^{p}}{\partial q_0^3}(\mathbf{0}) - 3c_3\frac{\partial^2\Theta^{nc}}{\partial q_0^2}(\mathbf{0}) - 3c_3^{+}\frac{\partial^2\Theta^{p}}{\partial q_0^2}(\mathbf{0})\\
B_{2,c}^{(3)} = c_2\frac{\partial^3\Theta^{nc}}{\partial q_0^3}(\mathbf{0}) + c_2^{+}\frac{\partial^3\Theta^{p}}{\partial q_0^3}(\mathbf{0}) - 3c_2^{-}\frac{\partial^2\Theta^{nc}}{\partial q_0^2}(\mathbf{0}) - 3c_2^{+}\frac{\partial^2\Theta^{p}}{\partial q_0^2}(\mathbf{0}),
\end{cases}
$$

$$
\begin{aligned}
B_{1,s}^{(3)} &= \begin{bmatrix} 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_1}(\mathbf{0}) & 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{3,s}^{(3)} &= \begin{bmatrix} 3c_3\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_1}(\mathbf{0}) & 3c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{2,s}^{(3)} &= \begin{bmatrix} 3c_2\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_1}(\mathbf{0}) & 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',
\end{aligned}
$$

$$
\begin{aligned}
B_{1qc} &= \begin{bmatrix} 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1^2}(\mathbf{0}) & 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & \mathbf{0}\\ 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_3^2}(\mathbf{0}) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix},\\
B_{3qc} &= \begin{bmatrix} 3c_3\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1^2}(\mathbf{0}) & 3c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & \mathbf{0}\\ 3c_3\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & 3c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_3^2}(\mathbf{0}) + c_3\frac{\partial^3\Theta^{ns}}{\partial q_3^2}(\mathbf{0}) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix},\\
B_{2qc} &= \begin{bmatrix} 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1^2}(\mathbf{0}) & 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & \mathbf{0}\\ 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_1\partial q_3}(\mathbf{0}) & 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_0\partial q_3^2}(\mathbf{0}) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
B_{1qs}^{(1)} &= \begin{bmatrix} c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_1^3}(\mathbf{0}) & 3c_1^{-}\frac{\partial^3\Theta^{nc}}{\partial q_1^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{1qs}^{(3)} &= \begin{bmatrix} 3c_1\frac{\partial^3\Theta^{nc}}{\partial q_1\partial q_3^2}(\mathbf{0}) & c_1^{-}\left(\frac{\partial^3\Theta^{nc}}{\partial q_3^3}(\mathbf{0}) + \frac{\partial^3\Theta^{ns}}{\partial q_3^3}(\mathbf{0})\right) & \mathbf{0}\end{bmatrix}',\\
B_{3qs}^{(1)} &= \begin{bmatrix} c_3\frac{\partial^3\Theta^{nc}}{\partial q_1^3}(\mathbf{0}) & 3c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_1^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{3qs}^{(3)} &= \begin{bmatrix} 3c_3\frac{\partial^3\Theta^{nc}}{\partial q_1\partial q_3^2}(\mathbf{0}) & c_3^{-}\frac{\partial^3\Theta^{nc}}{\partial q_3^3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{2qs}^{(1)} &= \begin{bmatrix} c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_1^3}(\mathbf{0}) & 3c_2^{-}\frac{\partial^3\Theta^{nc}}{\partial q_1^2\partial q_3}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',\\
B_{2qs}^{(3)} &= \begin{bmatrix} 3c_2\frac{\partial^3\Theta^{nc}}{\partial q_1\partial q_3^2}(\mathbf{0}) & c_2^{-}\left(\frac{\partial^3\Theta^{nc}}{\partial q_3^3}(\mathbf{0}) + \frac{\partial^3\Theta^{ns}}{\partial q_3^3}(\mathbf{0})\right) & \mathbf{0}\end{bmatrix}'.
\end{aligned}
$$

We also have

$$B_{1hc} = \begin{bmatrix} 3\sigma_1\rho_1 + 3c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 3c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix}, \quad B_{3hc} = \begin{bmatrix} 3c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 3c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) + c_3 \frac{\partial^2 \Theta^{ni}}{\partial q_3^2}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix}, \quad B_{2hc} = \begin{bmatrix} 3c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 3c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) \\ 3\sigma_2\rho_2 \end{bmatrix}$$

$$B_{1hs} = \begin{bmatrix} 3\sigma_1^2 + 3c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_1 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 3c_1 \left(\frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \Theta^{ni}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{3hs} = \begin{bmatrix} 3c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_3 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & c_3 \left(3\frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \Theta^{ni}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{2hs} = \begin{bmatrix} 3c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_2 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 3c_2 \left(\frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \Theta^{ni}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 3\sigma_2^2 \end{bmatrix}.$$

We further establish that

$$\frac{\partial \left[\frac{\partial^3 \alpha}{\partial u^3}(\mathbf{0}, \tau) \right]}{\partial \tau} = \alpha^{(3)}(\tau) + A_\alpha' \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau), \tag{35}$$

with

$$\begin{aligned}
\alpha^{(3)}(\tau) &= \alpha_c^{(3)} + \alpha_s^{(3)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau)' \alpha_q^{(3)}(\tau) \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + \alpha_h^{(3)}(\tau)' \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau), \\
\alpha_q^{(3)}(\tau) &= \alpha_{qc} + \alpha_{qs}(\tau), \quad \alpha_{qs}(\tau) = \begin{bmatrix} \alpha_{qs}^{(1)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \alpha_{qs}^{(3)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \alpha_{qs}^{(2)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) \end{bmatrix}, \\
\alpha_h^{(3)}(\tau) &= \alpha_{hc} + \alpha_{hs} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau), \quad for \quad j = 1, 2, 3,
\end{aligned}$$

and

$$\alpha_{qs}^{(2)\prime} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_c^{(3)} = c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0^3}(\mathbf{0}) + c_0^! \frac{\partial^3 \Theta^p}{\partial q_0^3}(\mathbf{0}) - 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_0^2}(\mathbf{0}) - 3c_0^! \frac{\partial^2 \Theta^p}{\partial q_0^2}(\mathbf{0}),$$

$$\alpha_s^{(3)} = \begin{bmatrix} 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0^2 \partial q_1}(\mathbf{0}) & 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$\alpha_{qc} = \begin{bmatrix} 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0 \partial q_1^2}(\mathbf{0}) & 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0 \partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0 \partial q_1 \partial q_3}(\mathbf{0}) & 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_0 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{qs}^{(1)} = \begin{bmatrix} c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad \alpha_{qs}^{(3)} = \begin{bmatrix} 3c_0 \frac{\partial^3 \Theta^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & c_0 \left(\frac{\partial^3 \Theta^{nc}}{\partial q_3^3}(\mathbf{0}) + \frac{\partial^3 \Theta^{ni}}{\partial q_3^3}(\mathbf{0}) \right) & \mathbf{0} \end{bmatrix}',$$

$$\alpha_{hc} = \begin{bmatrix} 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix},$$

$$\alpha_{hs} = \begin{bmatrix} 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_0 \frac{\partial^2 \Theta^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 3c_0 \left(\frac{\partial^2 \Theta^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \Theta^{ns}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Hence,

$$\begin{aligned} \frac{\partial^3 \alpha}{\partial u^3}(\mathbf{0}, \tau) &= \int_0^\tau \left(\alpha^{(3)}(s) + A'_\alpha \frac{\partial^3 \beta}{\partial u^3}(\mathbf{0}, s) \right) ds \\ &= \int_0^\tau \alpha^{(3)}(s)\, ds + A'_\alpha \int_0^\tau \frac{\partial^3 \beta}{\partial u^3}(\mathbf{0}, s)\, ds. \end{aligned} \tag{36}$$

Appendix B-4: Fourth cumulant

Similarly, the factor loading vector for the fourth cumulant is the solution of the following ODE:

$$\frac{\partial \left[\frac{\partial^4 \beta}{\partial u^4}(\mathbf{0}, \tau) \right]}{\partial \tau} = B^{(4)}(\tau) + A \frac{\partial^4 \beta}{\partial u^4}(\mathbf{0}, \tau) \tag{37}$$

with

$$B^{(4)}(\tau) = \left(B_1^{(4)}(\tau), B_3^{(4)}(\tau), B_2^{(4)}(\tau) \right)',$$

$$\begin{aligned} B_j^{(4)}(\tau) &= B_{j,c}^{(4)} + B_{j,s}^{(4)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau)' B_{j,q}^{(4)}(\tau) \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) \\ &\quad + B_{j,h}^{(4)}(\tau)' \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) + \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau)' B_{jhq}^{(4)} \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) + B_{j,l}^{(4)}(\tau)' \frac{\partial^3 \beta}{\partial u^3}(\mathbf{0}, \tau), \\ B_{j,q}^{(4)}(\tau) &= B_{jqc}^{(4)} + B_{jqs}^{(4)}(\tau), \end{aligned}$$

$$B_{jqs}^{(4)}(\tau) = \begin{bmatrix} B_{jqs}^{(4,1)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau)' B_{jqq}^{(4,1)} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + B_{jqsh}^{(4,1)\prime} \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) & B_{jqsh}^{(4,1,2)\prime} \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) & \mathbf{0} \\ B_{jqsh}^{(4,1,2)\prime} \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) & B_{jqs}^{(4,3)\prime} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau)' B_{jqq}^{(4,3)} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) + B_{jqsh}^{(4,3)\prime} \frac{\partial^2 \beta}{\partial u^2}(\mathbf{0}, \tau) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$B_{jk}^{(4)}(\tau) = B_{jkc}^{(4)} + B_{jks}^{(4)} \frac{\partial \beta}{\partial u}(\mathbf{0}, \tau) \quad for \quad j = 1, 2, 3, \quad and \quad k = h, l.$$

Moreover,

$$B_{jqs}^{(4,2)\prime} = B_{jqsh}^{(4,2)\prime} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix},$$

$$B_{jc}^{(4)} = c_j \frac{\partial^1 \Theta^{nc}}{\partial q_0^1}(\mathbf{0}) + c_j^1 \frac{\partial^1 \Theta^p}{\partial q_0^4}(\mathbf{0}) - 4c_j \frac{\partial^3 \Theta^{nc}}{\partial q_0^3}(\mathbf{0}) - 4c_j^1 \frac{\partial^3 \Theta^p}{\partial q_0^3}(\mathbf{0}),$$

$$B_{j,s}^{(4)} = \begin{bmatrix} 4c_j \frac{\partial^4 \Theta^{nc}}{\partial q_0^3 \partial q_1}(\mathbf{0}) & 4c_j \frac{\partial^4 \Theta^{nc}}{\partial q_0^3 \partial q_3}(\mathbf{0}) & 0 \end{bmatrix}'$$

$$B^{(4)}_{jqc} \quad \begin{bmatrix} 6c_j \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_1^2}(\mathbf{0}) & 6c_j \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 6c_j \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_1 \partial q_3}(\mathbf{0}) & 6c_j \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad j = 1, 2, 3,$$

$$B^{(4,1)}_{1qs} \quad \begin{bmatrix} 4c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^3}(\mathbf{0}) & 12c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{1qs} \quad \begin{bmatrix} 12c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1 \partial q_3^2}(\mathbf{0}) & 4c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1)}_{3qs} \quad \begin{bmatrix} 4c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^3}(\mathbf{0}) & 12c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{3qs} \quad \begin{bmatrix} 12c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1 \partial q_3^2}(\mathbf{0}) & 4c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_3^3}(\mathbf{0}) + c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1)}_{2qs} \quad \begin{bmatrix} 4c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^3}(\mathbf{0}) & 12c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{2qs} \quad \begin{bmatrix} 12c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_1 \partial q_3^2}(\mathbf{0}) & 4c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_0 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1)}_{1qq} \quad \begin{bmatrix} c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad B^{(4,3)}_{1qq} \quad \begin{bmatrix} 3c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_1 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_1 \left(\frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) + \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B^{(4,1)}_{3qq} \quad \begin{bmatrix} c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad B^{(4,3)}_{1qq} \quad \begin{bmatrix} 3c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_3 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B^{(4,1)}_{2qq} \quad \begin{bmatrix} c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad B^{(4,3)}_{2qq} \quad \begin{bmatrix} 3c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_2 \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_2 \left(\frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) + \frac{\partial^4 \leftrightarrow^{nc}}{\partial q_3^4}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B^{(4,1)}_{1qsh} \quad \begin{bmatrix} 6c_1 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^3}(\mathbf{0}) & 6c_1 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{1qsh} \quad \begin{bmatrix} 6c_1 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & 6c_1 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) + 6c_1 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1)}_{3qsh} \quad \begin{bmatrix} 6c_3 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^3}(\mathbf{0}) & 6c_3 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{3qsh} \quad \begin{bmatrix} 6c_3 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & 6c_3 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) + c_3 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1)}_{2qsh} \quad \begin{bmatrix} 6c_2 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^3}(\mathbf{0}) & 6c_2 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad B^{(4,3)}_{2qsh} \quad \begin{bmatrix} 6c_2 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & 6c_2 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) + 6c_2 \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4,1,2)'}_{jqsh} \quad \begin{bmatrix} 6c_j \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & 6c_j \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$B^{(4)}_{jhc} \quad \begin{bmatrix} 6c_j \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_1}(\mathbf{0}) \\ 6c_j \frac{\partial^3 \leftrightarrow^{nc}}{\partial q_0^2 \partial q_3}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix},$$

$$B_{1hs}^{(4)} \quad \begin{bmatrix} 12c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1^2}(\mathbf{0}) & 12c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 12c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & 12c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{3hs}^{(4)} \quad \begin{bmatrix} 12c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1^2}(\mathbf{0}) & 12c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 12c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & 12c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2^2}(\mathbf{0}) + 3c_3 \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{2hs}^{(4)} \quad \begin{bmatrix} 12c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1^2}(\mathbf{0}) & 12c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 12c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1 \partial q_2}(\mathbf{0}) & 12c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

$$B_{1lc}^{(4)} \quad \begin{bmatrix} 4\sigma_1 \rho_1 + 4c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 4c_1 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix}, \quad B_{3lc}^{(4)} \quad \begin{bmatrix} 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2}(\mathbf{0}) + c_3 \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix}, \quad B_{2lc}^{(4)} \quad \begin{bmatrix} 4c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 4c_2 \frac{\partial^2 \psi^{nc}}{\partial q_0 \partial q_2}(\mathbf{0}) \\ 4\sigma_2 \rho_2 \end{bmatrix},$$

$$B_{1ls}^{(4)} \quad \begin{bmatrix} 4\sigma_1^2 + 4c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 4c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 4c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 4c_1 \left(\frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{3ls}^{(4)} \quad \begin{bmatrix} 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 4c_3 \frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + 2c_3 \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{2ls}^{(4)} \quad \begin{bmatrix} 4c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 4c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 4c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 4c_2 \left(\frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 4\sigma_2^2 \end{bmatrix},$$

$$B_{1lq}^{(4)} \quad \begin{bmatrix} 3\sigma_1^2 + 3c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 3c_1 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 3c_1 \left(\frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{3lq}^{(4)} \quad \begin{bmatrix} 3c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 3c_3 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 3c_3 \frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + c_3 \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$B_{2lq}^{(4)} \quad \begin{bmatrix} 3c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & \mathbf{0} \\ 3c_2 \frac{\partial^2 \psi^{nc}}{\partial q_1 \partial q_2}(\mathbf{0}) & 3c_2 \left(\frac{\partial^2 \psi^{nc}}{\partial q_2^2}(\mathbf{0}) + \frac{\partial^2 \psi^{ns}}{\partial q_2^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 3\sigma_2^2 \end{bmatrix}.$$

Hence,

$$\frac{\partial^4 \beta}{\partial a^4}(0, \tau) = e^{\tau A} \left(\int_0^\tau e^{-sA} B^{(4)}(s)\, ds \right). \tag{38}$$

We further prove that

$$\frac{\partial\left[\frac{\partial^2\alpha}{\partial u^2}(\mathbf{0},\tau)\right]}{\partial t_\tau} = \alpha^{(4)}(\tau) + A_\alpha^{'}\frac{\partial^4\beta}{\partial u^4}(\mathbf{0},\tau), \tag{39}$$

with

$$\begin{aligned}\alpha^{(4)}(\tau) &= \alpha_c^{(4)} + \alpha_u^{(4)'}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \frac{\partial\beta}{\partial u}(\mathbf{0},\tau)'\alpha_q^{(4)}(\tau)\frac{\partial\beta}{\partial u}(\mathbf{0},\tau)\\ &\quad + \alpha_h^{(4)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) + \frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau)'\alpha_{hq}^{(4)}\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) + \alpha_l^{(4)}(\tau)'\frac{\partial^3\beta}{\partial u^3}(\mathbf{0},\tau),\\ \alpha_q^{(4)}(\tau) &= \alpha_{qc}^{(4)} + \alpha_{qu}^{(4)}(\tau),\end{aligned}$$

$$\alpha_{qu}^{(4)}(\tau) = \begin{bmatrix} \alpha_{qu}^{(4,1)'}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \frac{\partial\beta}{\partial u}(\mathbf{0},\tau)'\alpha_{qq}^{(4,1)}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \alpha_{quh}^{(4,1)'}\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) & \alpha_{quh}^{(4,2)'}\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) & \mathbf{0}\\ \alpha_{quh}^{(4,1,2)'}\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) & \alpha_{qu}^{(4,3)'}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \frac{\partial\beta}{\partial u}(\mathbf{0},\tau)'\alpha_{qq}^{(4,3)}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) + \alpha_{quh}^{(4,3)'}\frac{\partial^2\beta}{\partial u^2}(\mathbf{0},\tau) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix},$$

$$\alpha_k^{(4)}(\tau) = \alpha_{kc}^{(4)} + \alpha_{ku}^{(4)}\frac{\partial\beta}{\partial u}(\mathbf{0},\tau) \quad for \quad j = 1,2,3, \quad and \quad k = h,l.$$

We also define

$$\alpha_c^{(4)} = c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^4}(\mathbf{0}) + c_{(j)}^{\perp}\frac{\partial^4\phi^{p}}{\partial q_{(j)}^4}(\mathbf{0}) - 4c_{(j)}\frac{\partial^3\phi^{nc}}{\partial q_{(j)}^3}(\mathbf{0}) - 4c_{(j)}^{\perp}\frac{\partial^3\phi^{p}}{\partial q_{(j)}^3}(\mathbf{0}),$$

$$\alpha_u^{(4)} = \begin{bmatrix} 4c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^3\partial q_1}(\mathbf{0}) & 4c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^3\partial q_2}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',$$

$$\alpha_{qc}^{(4)} = \begin{bmatrix} 6c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^2\partial q_1^2}(\mathbf{0}) & 6c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^2\partial q_1\partial q_2}(\mathbf{0}) & \mathbf{0}\\ 6c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^2\partial q_1\partial q_2}(\mathbf{0}) & 6c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}^2\partial q_2^2}(\mathbf{0}) & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix},$$

$$\alpha_{qu}^{(4,1)} = \begin{bmatrix} 4c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}\partial q_1^2}(\mathbf{0}) & 12c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}\partial q_1^2\partial q_2}(\mathbf{0}) & \mathbf{0}\end{bmatrix}', \quad \alpha_{qu}^{(4,2)} = \begin{bmatrix} 12c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}\partial q_1\partial q_2^2}(\mathbf{0}) & 4c_{(j)}\frac{\partial^4\phi^{nc}}{\partial q_{(j)}\partial q_2^2}(\mathbf{0}) & \mathbf{0}\end{bmatrix}',$$

$$\alpha_{qq}^{(4,1)} = \begin{bmatrix} c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \alpha_{qq}^{(4,3)} = \begin{bmatrix} 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_0^- \left(\frac{\partial^4 \ominus^{nc}}{\partial q_3^4}(\mathbf{0}) + \frac{\partial^4 \ominus^{nt}}{\partial q_3^4}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{qq}^{(4,1)} = \begin{bmatrix} c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \alpha_{qq}^{(4,3)} = \begin{bmatrix} 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_3^4}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{qq}^{(4,1)} = \begin{bmatrix} c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^4}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^3 \partial q_3}(\mathbf{0}) & 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \alpha_{qq}^{(4,3)} = \begin{bmatrix} 3c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1^2 \partial q_3^2}(\mathbf{0}) & 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & \mathbf{0} \\ 2c_0^- \frac{\partial^4 \ominus^{nc}}{\partial q_1 \partial q_3^3}(\mathbf{0}) & c_0^- \left(\frac{\partial^4 \ominus^{nc}}{\partial q_3^4}(\mathbf{0}) + \frac{\partial^4 \ominus^{nt}}{\partial q_3^4}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{qsh}^{(4,1)} = \begin{bmatrix} 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_1^3}(\mathbf{0}) & 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}', \quad \alpha_{qsh}^{(4,3)} = \begin{bmatrix} 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_3^3}(\mathbf{0}) + 6c_0^- \frac{\partial^3 \ominus^{nt}}{\partial q_3^3}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$\alpha_{qsh}^{(4,1,2)'} = \begin{bmatrix} 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_1^2 \partial q_3}(\mathbf{0}) & 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_1 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \end{bmatrix}',$$

$$\alpha_{hc}^{(4)} = \begin{bmatrix} 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0^2 \partial q_1}(\mathbf{0}) \\ 6c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0^2 \partial q_3}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix},$$

$$\alpha_{hs}^{(4)} = \begin{bmatrix} 12c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0 \partial q_1^2}(\mathbf{0}) & 12c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0 \partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 12c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0 \partial q_1 \partial q_3}(\mathbf{0}) & 12c_0^- \frac{\partial^3 \ominus^{nc}}{\partial q_0 \partial q_3^2}(\mathbf{0}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{fc}^{(4)} = \begin{bmatrix} 4c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_0 \partial q_1}(\mathbf{0}) \\ 4c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_0 \partial q_3}(\mathbf{0}) \\ \mathbf{0} \end{bmatrix},$$

$$\alpha_{fs}^{(4)} = \begin{bmatrix} 4c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1^2}(\mathbf{0}) & 4c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 4c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 4c_0^- \left(\frac{\partial^2 \ominus^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \ominus^{nt}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\alpha_{hq}^{(4)} = \begin{bmatrix} 3c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1^2}(\mathbf{0}) & 3c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & \mathbf{0} \\ 3c_0^- \frac{\partial^2 \ominus^{nc}}{\partial q_1 \partial q_3}(\mathbf{0}) & 3c_0^- \left(\frac{\partial^2 \ominus^{nc}}{\partial q_3^2}(\mathbf{0}) + \frac{\partial^2 \ominus^{nt}}{\partial q_3^2}(\mathbf{0}) \right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

Hence,

$$\begin{aligned}\frac{\partial^4 \alpha}{\partial u^4}(0,\tau) &= \int_0^\tau \left(\alpha^{(4)}(s) + A'_\alpha \frac{\partial^4 \beta}{\partial u^4}(0,s)\right) ds \\ &= \int_0^\tau \alpha^{(4)}(s)\, ds + A'_\alpha \int_0^\tau \frac{\partial^4 \beta}{\partial u^4}(0,s)\, ds. \end{aligned} \tag{40}$$

Appendix C: General risk-neutral moments ODEs

We provide the general expressions of $\alpha^{(n)}(\tau)$ and $B^{(n)}(\tau) = \left(B_j^{(n)}(\tau)\right)_{j=1,..N}$ in the ODEs for risk-neutral moments. We have

$$\begin{aligned} B_j^{(2)}(\tau) &= B_{j,c}^{(2)} + B_{j,s}^{(2)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)' B_{j,q}^{(2)}\frac{\partial\beta}{\partial u}(0,\tau), \\ \alpha^{(2)}(\tau) &= \alpha_c^{(2)} + \alpha_s^{(2)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)' \alpha_q^{(2)}\frac{\partial\beta}{\partial u}(0,\tau). \end{aligned}$$

Moreover,

$$\begin{aligned} B_j^{(3)}(\tau) &= B_{j,c}^{(3)} + B_{j,s}^{(3)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)' B_{j,q}^{(3)}(\tau)\frac{\partial\beta}{\partial u}(0,\tau) + B_{j,h}^{(3)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(0,\tau), \\ B_{j,q}^{(3)}(\tau) &= B_{jqc} + B_{jqs}(\tau), \quad B_{jqs}(\tau) = diag\left(B_{jqs}^{(i)\prime}\frac{\partial\beta}{\partial u}(0,\tau), i = 1,...,N\right), \\ B_{j,h}^{(3)}(\tau) &= B_{jhc} + B_{jhs}\frac{\partial\beta}{\partial u}(0,\tau), \end{aligned}$$

$$\begin{aligned} \alpha^{(3)}(\tau) &= \alpha_c^{(3)} + \alpha_s^{(3)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)'\alpha_q^{(3)}(\tau)\frac{\partial\beta}{\partial u}(0,\tau) + \alpha_h^{(3)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(0,\tau), \\ \alpha_q^{(3)}(\tau) &= \alpha_{qc} + \alpha_{qs}(\tau), \quad with \quad \alpha_{qs}(\tau) = diag\left(\alpha_{qs}^{(i)\prime}\frac{\partial\beta}{\partial u}(0,\tau), i = 1,...,N\right), \\ \alpha_h^{(3)}(\tau) &= \alpha_{hc} + \alpha_{hs}\frac{\partial\beta}{\partial u}(0,\tau). \end{aligned}$$

In addition,

$$\begin{aligned} B_j^{(4)}(\tau) &= B_{j,c}^{(4)} + B_{j,s}^{(4)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)' B_{j,q}^{(4)}(\tau)\frac{\partial\beta}{\partial u}(0,\tau) \\ &\quad + B_{j,h}^{(4)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(0,\tau) + \frac{\partial^2\beta}{\partial u^2}(0,\tau)' B_{jhq}^{(4)}\frac{\partial^2\beta}{\partial u^2}(0,\tau) + B_{j,l}^{(4)}(\tau)'\frac{\partial^3\beta}{\partial u^3}(0,\tau), \\ B_{j,q}^{(4)}(\tau) &= B_{jqc}^{(4)} + B_{jqs}^{(4)}(\tau), \quad with \end{aligned}$$

$$\begin{aligned} B_{jqs}^{(4)}(\tau) &= diag\left(B_{jqs}^{(4,i)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)' B_{jqq}^{(4,i)}\frac{\partial\beta}{\partial u}(0,\tau), i = 1,...,N\right) \\ &\quad + \left(B_{jqsh}^{(4,i,k)\prime}\frac{\partial^2\beta}{\partial u^2}(0,\tau)\right)_{i,k=1,...,N}, \end{aligned}$$

$$B_{jk}^{(4)}(\tau) = B_{jkc}^{(4)} + B_{jks}^{(4)}\frac{\partial\beta}{\partial u}(0,\tau), \quad for \quad j = 1,...,N, \quad k = h,l$$

$$\begin{aligned} \alpha^{(4)}(\tau) &= \alpha_c^{(4)} + \alpha_s^{(4)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)'\alpha_q^{(4)}(\tau)\frac{\partial\beta}{\partial u}(0,\tau) \\ &\quad + \alpha_h^{(4)}(\tau)'\frac{\partial^2\beta}{\partial u^2}(0,\tau) + \frac{\partial^2\beta}{\partial u^2}(0,\tau)'\alpha_{hq}^{(4)}\frac{\partial^2\beta}{\partial u^2}(0,\tau) + \alpha_l^{(4)}(\tau)'\frac{\partial^3\beta}{\partial u^3}(0,\tau), \\ \alpha_q^{(4)}(\tau) &= \alpha_{qc}^{(4)} + \alpha_{qs}^{(4)}(\tau), \quad with \end{aligned}$$

$$\alpha_{qs}^{(4)}(\tau) = diag\left(\alpha_{qs}^{(4,i)\prime}\frac{\partial\beta}{\partial u}(0,\tau) + \frac{\partial\beta}{\partial u}(0,\tau)'\alpha_{qq}^{(4,i)}\frac{\partial\beta}{\partial u}(0,\tau) \ , i=1,...,N\right)$$
$$+\left(\alpha_{qsh}^{(4,i,k)\prime}\frac{\partial^2\beta}{\partial u^2}(0,\tau)\right)_{i,k=1,...,N},$$

where $B_{j,q}^{(2)}$, $\alpha_q^{(2)}$, B_{jhc}, α_{qc}, $B_{jqc}^{(4)}$, $B_{jhq}^{(4)}$, $\alpha_{qc}^{(4)}$ and $\alpha_{hq}^{(4)}$ are symmatric matrices and $B_{jqsh}^{(4,i,k)} = B_{jqsh}^{(4,k,i)}$, $\alpha_{qsh}^{(4,i,k)} = \alpha_{qsh}^{(4,k,i)}$.

Appendix D: Details on the AFT model estimation

Appendix D-1: Discretizing the stochastic jump-diffusion motion

Starting from the underlying price dynamics in Andersen et al. (2015b) model, we compute the quadratic variation of $V_{1t}^J \equiv \int_0^t \int_{R^2} x^2 1_{\{x<0\}} \mu(dt,dx,dy)$ and $V_{3t}^J \equiv \int_0^t \int_{R^2} \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right] \mu(dt,dx,dy)$, as well as their covariation. We have

$$\begin{aligned}
\left[V_1^J, V_1^J\right]_t &= \int_0^t \int_{R^2} x^4 1_{\{x<0\}} \mu(dt,dx,dy), \quad \left[V_3^J, V_3^J\right]_t = \int_0^t \int_{R^2} \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right]^2 \mu(dt,dx,dy), \\
\left[V_1^J, V_3^J\right]_t &= \int_0^t \int_{R^2} x^2 \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right] 1_{\{x<0\}} \mu(dt,dx,dy).
\end{aligned}$$

We also have the corresponding risk-neutral expectations

$$\begin{aligned}
E^Q\left[V_1^J, V_1^J\right]_t &= \int_0^t \int_{R^2} x^4 1_{\{x<0\}} \nu_t^Q(dx,dy)\, dt, \quad E^Q\left[V_3^J, V_3^J\right]_t = \int_0^t \int_{R^2} \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right]^2 \nu_t^Q(dx,dy)\, dt, \\
E^Q\left[V_1^J, V_3^J\right]_t &= \int_0^t \int_{R^2} x^2 \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right] 1_{\{x<0\}} \nu_t^Q(dx,dy)\, dt.
\end{aligned}$$

Thus, the instantaneous variance and covariance of V_{1t}^J and V_{3t}^J can be computed as

$$\begin{aligned}
\left(\sigma_{1t}^J\right)^2 &= \int_{R^2} x^4 1_{\{x<0\}} \nu_t^Q(dx,dy), \\
\left(\sigma_{3t}^J\right)^2 &= \int_{R^2} \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right]^2 \nu_t^Q(dx,dy), \\
\sigma_{13t}^J &= \int_{R^2} x^2 \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right] 1_{\{x<0\}} \nu_t^Q(dx,dy).
\end{aligned}$$

Formally, the instantaneous variance is

$$\begin{aligned}
\left(\sigma_{1t}^J\right)^2 &= \int_{R^2} x^4 1_{\{x<0\}} \nu_t^Q(dx,dy), \\
&= c^- \int_{R^2} 1_{\{x<0, y=0\}} x^4 \lambda_- e^{-\lambda_-|x|} dx \otimes dy, \\
&= c^- \int_{-\infty}^0 x^4 \lambda_- e^{-\lambda_-|x|} dx \equiv \lambda^* c_t^-,
\end{aligned} \tag{41}$$

with

$$\lambda^* = \int_{-\infty}^0 x^4 \lambda_- e^{-\lambda_-|x|} dx = \frac{24}{\lambda^4}.$$

Similarly,

$$\begin{aligned}
\left(\sigma_{3t}^J\right)^2 &= \int_{R^2} \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right]^2 \nu_t^Q(dx,dy), \\
&= \int_{R^2} \left[(1-\rho_3)^2 x^4 1_{\{x<0\}} + \rho_3^2 y^4 + 2\rho_3 (1-\rho_3) x^2 y^2 1_{\{x<0\}}\right] \nu_t^Q(dx,dy), \\
&= \int_{R^2} \left[\left((1-\rho_3)^2 x^4 + 2\rho_3 (1-\rho_3) x^2 y^2\right) 1_{\{x<0\}} \nu_t^Q(dx,dy) + \rho_3^2 y^4 \nu_t^Q(dx,dy)\right], \\
&= \int_{R^2} \left((1-\rho_3)^2 x^4 + 2\rho_3 (1-\rho_3) x^2 y^2\right) 1_{\{x<0\}} \nu_t^Q(dx,dy) + \int_{R^2} \rho_3^2 y^4 \nu_t^Q(dx,dy).
\end{aligned}$$

Note that

$$\begin{aligned}&\left((1-\rho_3)^2 x^4 + 2\rho_3 (1-\rho_3) x^2 y^2\right) 1_{\{x<0\}} \nu_t^Q (dx, dy)\\ =\ & c^- \left((1-\rho_3)^2 x^4 + 2\rho_3 (1-\rho_3) x^2 y^2\right) \lambda_- e^{-\lambda_- |x|} 1_{\{x<0, y=0\}} dx \otimes dy,\end{aligned}$$

and

$$\begin{aligned}&\int_{R^2} \left((1-\rho_3)^2 x^4 + 2\rho_3 (1-\rho_3) x^2 y^2\right) 1_{\{x<0\}} \nu_t^Q (dx, dy)\\ =\ & c^- (1-\rho_3)^2 \int_{-\infty}^{0} x^4 \lambda_- e^{-\lambda_- |x|} dx = (1-\rho_3)^2 \lambda_-^4 c_t^-.\end{aligned}$$

Moreover,

$$\rho_3^2 y^4 \nu_t^Q (dx, dy) = \rho_3^2 y^4 \left\{ \left(c^- 1_{\{x<0\}} \lambda_- e^{-\lambda_- |x|} + c^+ 1_{\{x>0\}} \lambda_+ e^{-\lambda_+ |x|} \right) 1_{\{y=0\}} + c^- 1_{\{x=0, y<0\}} \lambda_- e^{-\lambda_- |y|} \right\} dx \otimes dy,$$

and

$$\int_{R^2} \rho_3^2 y^4 \nu_t^Q (dx, dy) = \rho_3^2 \left[\int_{-\infty}^{0} y^4 \lambda_- e^{-\lambda_- |y|} dy \right] c_t^- = \rho_3^2 \lambda_-^4 c_t^-.$$

This yields

$$\left(\sigma_{3t}^J\right)^2 = \left[(1-\rho_3)^2 + \rho_3^2\right] \lambda_-^4 c_t^-. \tag{42}$$

For the instantaneous covariance, we have

$$\begin{aligned}\sigma_{13t}^J &= \int_{R^2} x^2 \left[(1-\rho_3) x^2 1_{\{x<0\}} + \rho_3 y^2\right] 1_{\{x<0\}} \nu_t^Q (dx, dy),\\ &= (1-\rho_3) \int_{R^2} x^4 1_{\{x<0\}} \nu_t^Q (dx, dy) + \rho_3 \int_{R^2} x^2 y^2 1_{\{x<0\}} \nu_t^Q (dx, dy),\\ &= (1-\rho_3) \left(\sigma_{1t}^J\right)^2.\end{aligned} \tag{43}$$

Turning to the risk-neutral expectations, we get

$$\begin{aligned}&E^Q \left[\mu_1 \int_{R^2} x^2 1_{\{x<0\}} \mu (dt, dx, dy) \right]\\ =\ & \mu_1 \left(\int_{R^2} x^2 1_{\{x<0\}} \nu_t^Q (dx, dy) \right) dt,\\ =\ & \mu_1 dt \left(\int_{-\infty}^{0} x^2 \lambda_- e^{-\lambda_- |x|} dx \right) c_t^-,\\ \equiv\ & \mu_1 dt \lambda_- c_t^-,\end{aligned}$$

with

$$\lambda_- = \int_{-\infty}^{0} x^2 \lambda_- e^{-\lambda_- |x|} dx = \frac{2}{\lambda_-^2}.$$

In the same way, we obtain

$$
\begin{aligned}
& E^{Q}\left[\mu_3 \int_{R^2}\left[(1-\rho_3)\, x^2 1_{\{x<0\}}+\rho_3 y^2\right]\mu\left(dt,dx,dy\right)\right] \\
= \;& \mu_3 dt \int_{R^2}\left[(1-\rho_3)\, x^2 1_{\{x<0\}}+\rho_3 y^2\right]\nu_t^{Q}\left(dx,dy\right), \\
= \;& \mu_3 dt \int_{R^2}\left[(1-\rho_3)\, x^2 1_{\{x<0\}}+\rho_3 y^2\right]\left\{\begin{array}{c}\left(c_t^{-} 1_{\{x<0\}}\lambda_{-}\, e^{-\lambda_{-}|x|}+c_t^{+} 1_{\{x>0\}}\lambda_{+}\, e^{-\lambda_{+}|x|}\right) 1_{\{y=0\}} \\ +c_t^{-} 1_{\{x=0,y<0\}}\lambda_{-}\, e^{-\lambda_{-}|y|}\end{array}\right\} dx\otimes dy, \\
= \;& \mu_3\left(1-\rho_3\right) dt\left(\int_{-\infty}^{0} x^2\lambda_{-}\, e^{-\lambda_{-}|x|}dx\right)c_t^{-}+\mu_3\rho_3 dt\left(\int_{-\infty}^{0} y^2\lambda_{-}\, e^{-\lambda_{-}|y|}dy\right)c_t^{-}, \\
= \;& \mu_3 dt\left(\int_{-\infty}^{0} x^2\lambda_{-}\, e^{-\lambda_{-}|x|}dx\right)c_t^{-}=\mu_3 dt\lambda_{-}\, c_t^{-}. \qquad (44)
\end{aligned}
$$

The resulting discretized transition equations can be written as

$$
\begin{aligned}
V_{1t+1}-V_{1t} &= \kappa_1\left(\bar{c}_1-V_{1t}\right)\Delta t+\mu_1\lambda_{-}\, c_t\,\Delta t+\varepsilon_{1t+1}, & (45)\\
V_{2t+1}-V_{2t} &= \kappa_2\left(\bar{c}_2-V_{2t}\right)\Delta t+\varepsilon_{2t+1}, & (46)\\
V_{3t+1}-V_{3t} &= -\kappa_3 V_{3t}\Delta t+\mu_3\lambda_{-}\, c_t^{-}\Delta t+\varepsilon_{3t+1}, & (47)
\end{aligned}
$$

or in a compact form as

$$V_{t+1}=\Phi_0+\Phi_1 V_t+\varepsilon_{t+1},$$

with

$$\Phi_0\equiv\Delta t\begin{pmatrix}\kappa_1\bar{c}_1+\mu_1\lambda_{-}\, c_0^{-}\\ \kappa_2\bar{c}_2\\ \mu_3\lambda_{-}\, c_0^{-}\end{pmatrix},\quad \Phi_1\equiv I_3+K_1,\quad K_1=\Delta t\begin{bmatrix}-\kappa_1+\mu_1\lambda_{-}\, c_1^{-} & \mu_1\lambda_{-}\, c_2^{-} & \mu_1\lambda_{-}\, c_3^{-}\\ 0 & -\kappa_2 & 0\\ \mu_3\lambda_{-}\, c_1^{-} & \mu_3\lambda_{-}\, c_2^{-} & -\kappa_3+\mu_3\lambda_{-}\, c_3^{-}\end{bmatrix}.$$

Specifically, I_3 is a 3×3 identity matrix, $V_{t+1}\equiv\left(V_{1t+1},V_{2t+1},V_{3t+1}\right)'$, and $\varepsilon_{t+1}\equiv\left(\varepsilon_{1t+1},\varepsilon_{2t+1},\varepsilon_{3t+1}\right)'$. The covariance matrix of the noise term is

$$Var_t\left(\varepsilon_{t+1}\right)=\Delta t\begin{bmatrix}\sigma_1^2 V_{1t}+\mu_1^2\lambda_{-}^{*}c_t^{-} & 0 & \mu_1\mu_3\left(1-\rho_3\right)\lambda_{-}^{*}c_t^{-}\\ 0 & \sigma_2^2 V_{2t} & 0\\ \mu_1\mu_3\left(1-\rho_3\right)\lambda_{-}^{*}c_t^{-} & 0 & \mu_3^2\left[\left(1-\rho_3\right)^2+\rho_3^2\right]\lambda_{-}^{*}c_t^{-}\end{bmatrix}.$$

At the daily frequency, the discrete time step is $\Delta t=1/252$. Combining the state transition equation with the risk-neutral cumulant measurement equation (outlined in the main text) allows to estimate the factors along with the parameters of the model using the modified Kalman filter algorithm. The initial conditions are $V_{0|0}=-K_1^{-1}\Phi_0$, and $vec\left(P_{0|0}\right)=\left(I_9-\Phi_1\otimes\Phi_1\right)^{-1}vec\left(\Sigma\left(V_{0|0}\right)\right)$, where $\otimes$ denotes the Kronecker product.

Appendix D-2: Parameter constraints

Our goal here is to check whether, beyond the parameter restrictions imposed by Andersen et al. (2015b), additional constraints are required in our estimation procedure to guarantee admissibility of the three latent factors and covariance stationarity.

Diagonalization The implementation of the risk-neutral moment-based estimation for the three-factor AFT model was done under the condition that the matrix K_1 is diagonalizable. This is

equivalent to ensuring that the block triangular matrix

$$K_1^s = \Delta t \begin{bmatrix} -\kappa_1 + \mu_1\lambda_- c_1^- & \mu_3\lambda_- c_1^- & 0 \\ \mu_1\lambda_- c_3^- & -\kappa_3 + \mu_3\lambda_- c_3^- & 0 \\ \mu_1\lambda_- c_2^- & \mu_3\lambda_- c_2^- & -\kappa_2 \end{bmatrix} \tag{48}$$

is diagonalizable. Note that K_1^s is obtained from K_1 by (i) permuting the second and third columns, (ii) then permuting the second and third rows, (iii) and finally transposing the resulting matrix. Recall that $\lambda_- = 2/\lambda_-^2$. Thus, verifying if K_1 is diagonalizable, boils down to checking if

$$\hat{A} = \begin{bmatrix} -\kappa_1 + \frac{2}{\lambda^2}\mu_1 c_1^- & \frac{2}{\lambda^2}\mu_3 c_1^- \\ \frac{2}{\lambda^2}\mu_1 c_3^- & -\kappa_3 + \frac{2}{\lambda^2}\mu_3 c_3^- \end{bmatrix}, \tag{49}$$

the upper left block in K_1^s, is diagonalizable. By computing

$$\begin{aligned} & \det\left(\hat{A} - \lambda I_2\right) \\ = & \left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^- - \lambda\right)\left(-\kappa_3 + \frac{2}{\lambda_-^2}\mu_3 c_3^- - \lambda\right) - \frac{2}{\lambda_-^2}\mu_1 c_3^- \frac{2}{\lambda_-^2}\mu_3 c_1^-, \\ = & \left(\lambda + \kappa_1 - \frac{2}{\lambda_-^2}\mu_1 c_1^-\right)\left(\lambda + \kappa_3 - \frac{2}{\lambda_-^2}\mu_3 c_3^-\right) - \frac{2}{\lambda_-^2}\frac{2}{\lambda_-^2}\mu_3 c_1^-\mu_1 c_3^-, \\ = & \lambda^2 - \left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^- - \kappa_3 + \frac{2}{\lambda_-^2}\mu_3 c_3^-\right)\lambda + \left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^-\right)\left(-\kappa_3 + \frac{2}{\lambda_-^2}\mu_3 c_3^-\right) - \frac{4}{\lambda_-^4}\mu_3 c_1^-\mu_1 c_3^-, \end{aligned}$$

we see that a sufficient condition for diagonalization is

$$\left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^- - \kappa_3 + \frac{2}{\lambda_-^2}\mu_3 c_3^-\right)^2 - 4\left(\left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^-\right)\left(-\kappa_3 + \frac{2}{\lambda_-^2}\mu_3 c_3^-\right) - \frac{4}{\lambda_-^4}\mu_3 c_1^-\mu_1 c_3^-\right) > 0.$$

This implies that

$$\left(-\kappa_1 + \frac{2}{\lambda_-^2}\mu_1 c_1^- + \kappa_3 - \frac{2}{\lambda_-^2}\mu_3 c_3^-\right)^2 + \frac{16}{\lambda_-^4}\mu_3 c_1^-\mu_1 c_3^- > 0, \tag{50}$$

which is always true. The corresponding eigenvalues are

$$\lambda_1 = \frac{\left(-\kappa_1 + \lambda_-\mu_1 c_1^- - \kappa_3 + \lambda_-\mu_3 c_3^-\right) + \sqrt{\left(-\kappa_1 + \lambda_-\mu_1 c_1^- + \kappa_3 - \lambda_-\mu_3 c_3^-\right)^2 + 4\left(\lambda^2\mu_3 c_1^-\mu_1 c_3^-\right)}}{2}, \tag{51}$$

$$\lambda_2 = \frac{\left(-\kappa_1 + \lambda_-\mu_1 c_1^- - \kappa_3 + \lambda_-\mu_3 c_3^-\right) - \sqrt{\left(-\kappa_1 + \lambda_-\mu_1 c_1^- + \kappa_3 - \lambda_-\mu_3 c_3^-\right)^2 + 4\left(\lambda^2\mu_3 c_1^-\mu_1 c_3^-\right)}}{2}. \tag{52}$$

For a given eigenvalue $\lambda \in \{\lambda_1, \lambda_2\}$, solving for the corresponding eigenvector $\hat{P} = \begin{bmatrix} \hat{P}_1 & \hat{P}_2 \end{bmatrix}^t$ in

$$\hat{A}\hat{P} = \lambda\hat{P},$$

or equivalently in

$$\begin{cases} \left(-\kappa_1 + \lambda_-\mu_1 c_1^- - \lambda\right)\hat{P}_1 + \lambda_-\mu_3 c_1^-\hat{P}_2 = 0, \\ \lambda_-\mu_1 c_3^-\hat{P}_1 + \left(-\kappa_3 + \lambda_-\mu_3 c_3^- - \lambda\right)\hat{P}_2 = 0, \end{cases}$$

yields

$$P_2 = \frac{\left(\lambda + \kappa_1 - \lambda_- \mu_1 c_1^-\right)}{\lambda_- \mu_3 c_1^-} P_1. \tag{53}$$

Hence, an eigenvector associated with $\lambda \in \{\lambda_1, \lambda_2\}$ is

$$\begin{pmatrix} \lambda_- \mu_3 c_1^- \\ \lambda + \kappa_1 - \lambda_- \mu_1 c_1^- \end{pmatrix}. \tag{54}$$

This confirms that the matrix K_1 is diagonalizable.

Stationarity To ensure covariance stationarity in the AFT model, the eigenvalues of K_1 should be between -2 and 0. Thus,

$$\det\left(K_1 - \lambda I_3\right) = (\Delta t)^3 \det\left(\begin{bmatrix} -\kappa_1 + \mu_1 \lambda_- c_1^- - \frac{\lambda}{\Delta t} & \mu_1 \lambda_- c_2^- & \mu_1 \lambda_- c_3^- \\ 0 & -\kappa_2 - \frac{\lambda}{\Delta t} & 0 \\ \mu_3 \lambda_- c_1^- & \mu_3 \lambda_- c_2^- & -\kappa_3 + \mu_3 \lambda_- c_3^- - \frac{\lambda}{\Delta t} \end{bmatrix}\right). \tag{55}$$

For the eigenvalue λ to be between -2 and 0, it must be the case that

$$\begin{aligned} -2 < \lambda < 0 &\iff -\frac{2}{\Delta t} < \frac{\lambda}{\Delta t} < 0 \\ &\iff -\frac{2}{\Delta t} < \hat{\lambda} < 0, \end{aligned}$$

where $\hat{\lambda} = \lambda / \Delta t$. Next, the determinant is computed as

$$\begin{aligned} & \det\left(\begin{bmatrix} -\kappa_1 + \mu_1 \lambda_- c_1^- - \hat{\lambda} & \mu_1 \lambda_- c_2^- & \mu_1 \lambda_- c_3^- \\ 0 & -\kappa_2 - \hat{\lambda} & 0 \\ \mu_3 \lambda_- c_1^- & \mu_3 \lambda_- c_2^- & -\kappa_3 + \mu_3 \lambda_- c_3^- - \hat{\lambda} \end{bmatrix}\right) \\ = & \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \hat{\lambda}\right)\left(-\kappa_2 - \hat{\lambda}\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^- - \hat{\lambda}\right) - \mu_1 \lambda_- c_3^- \left(-\kappa_2 - \hat{\lambda}\right) \mu_3 \lambda_- c_1^-, \\ = & \left(-\kappa_2 - \hat{\lambda}\right)\left(\left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \hat{\lambda}\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^- - \hat{\lambda}\right) - \mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-\right), \\ = & \left(-\kappa_2 - \hat{\lambda}\right)\left(\hat{\lambda}^2 - \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right)\hat{\lambda} + \left(-\kappa_1 + \mu_1 \lambda_- c_1^-\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^-\right) - \mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-\right), \\ = & \left(-\kappa_2 - \hat{\lambda}\right)\left(\hat{\lambda}^2 - \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right)\hat{\lambda} + \kappa_1 \kappa_3 - \kappa_1 \mu_3 \lambda_- c_3^- - \kappa_3 \mu_1 \lambda_- c_1^-\right). \end{aligned} \tag{56}$$

The discriminant of this second-order polynomial in $\hat{\lambda}$ is

$$\begin{aligned} \Delta &= \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right)^2 - 4\left(\kappa_1 \kappa_3 - \kappa_1 \mu_3 \lambda_- c_3^- - \kappa_3 \mu_1 \lambda_- c_1^-\right), \\ &= \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right)^2 - 4\left(\left(-\kappa_1 + \mu_1 \lambda_- c_1^-\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^-\right) - \mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-\right), \\ &= \left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right)^2 - 4\left(-\kappa_1 + \mu_1 \lambda_- c_1^-\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^-\right) + 4\mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-, \\ &= \left(-\kappa_1 + \mu_1 \lambda_- c_1^- + \kappa_3 - \mu_3 \lambda_- c_3^-\right)^2 + 4\mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^- > 0, \end{aligned} \tag{57}$$

and the roots are

$$\lambda_1 = \frac{\left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right) - \sqrt{\left(-\kappa_1 + \mu_1 \lambda_- c_1^- + \kappa_3 - \mu_3 \lambda_- c_3^-\right)^2 + 4\mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-}}{2}, \tag{58}$$

$$\lambda_2 = \frac{\left(-\kappa_1 + \mu_1 \lambda_- c_1^- - \kappa_3 + \mu_3 \lambda_- c_3^-\right) + \sqrt{\left(-\kappa_1 + \mu_1 \lambda_- c_1^- + \kappa_3 - \mu_3 \lambda_- c_3^-\right)^2 + 4\mu_1 \lambda_- c_3^- \mu_3 \lambda_- c_1^-}}{2}. \tag{59}$$

The stationarity condition translates into the following constraints:

$$\begin{cases} -\frac{2}{\Delta t} < -\kappa_2 < 0, \\ -\frac{2}{\Delta t} < \dot{\lambda}_1 < 0, \\ -\frac{2}{\Delta t} < \dot{\lambda}_2 < 0. \end{cases} \tag{60}$$

Hence,

$$0 < \kappa_2 < \frac{2}{\Delta t}, \tag{61}$$

and

$$\begin{cases} -\frac{2}{\Delta t} < \frac{\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) - \sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-}}{2} < 0, \\ -\frac{2}{\Delta t} < \frac{\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) + \sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-}}{2} < 0. \end{cases} \tag{62}$$

Alternatively, one can consider the following set of inequalities:

$$\begin{cases} 0 < \frac{\sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-} - \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right)}{2} < \frac{2}{\Delta t}, \\ 0 < \frac{-\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) - \sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-}}{2} < \frac{2}{\Delta t}, \end{cases}$$

which entails

$$\begin{aligned} &\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) < \\ &\sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-} \\ &< \frac{4}{\Delta t} + \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right), \end{aligned}$$

and

$$\begin{aligned} &-\frac{4}{\Delta t} - \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) \\ &< \sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-} \\ &< -\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right). \end{aligned} \tag{63}$$

If

$$\frac{4}{\Delta t} + 2\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) > 0, \tag{64}$$

then

$$\frac{4}{\Delta t} + \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) > -\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right),$$

and

$$-\frac{4}{\Delta t} - \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) < \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right),$$

which further implies

$$\begin{aligned} &\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) \\ < &\sqrt{\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^-} \\ < &-\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right). \end{aligned} \tag{65}$$

It follows that $\left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right) < 0$ and

$$\left(-\kappa_1 + \mu_1\lambda_- c_1^- + \kappa_3 - \mu_3\lambda_- c_3^-\right)^2 + 4\mu_1\lambda_- c_3^-\mu_3\lambda_- c_1^- < \left(-\kappa_1 + \mu_1\lambda_- c_1^- - \kappa_3 + \mu_3\lambda_- c_3^-\right)^2.$$

Hence,

$$4\mu_1\lambda_-c_3^-\mu_3\lambda_-c_1^- < 4\left(-\kappa_1 + \mu_1\lambda_-c_1^-\right)\left(-\kappa_3 + \mu_3\lambda_-c_3^-\right), \tag{66}$$

and

$$\frac{\mu_3\lambda_-c_3^-\kappa_1}{\kappa_1 - \mu_1\lambda_-c_1^-} < \kappa_3. \tag{67}$$

It must also be the case that

$$\kappa_1 - \mu_1\lambda_-c_1^- > 0, \tag{68}$$

or equivalently

$$\mu_1\lambda_-c_1^- < \kappa_1.$$

We now verify whether the condition $\left(-\kappa_1 + \mu_1\lambda_-c_1^- - \kappa_3 + \mu_3\lambda_-c_3^-\right) < 0$ is met. We have

$$\begin{aligned} & -\kappa_1 + \mu_1\lambda_-c_1^- - \kappa_3 + \mu_3\lambda_-c_3^- \\ < \; & -\kappa_1 + \mu_1\lambda_-c_1^- - \frac{\mu_3\lambda_-c_3^-\kappa_1}{\kappa_1 - \mu_1\lambda_-c_1^-} + \mu_3\lambda_-c_3^- \\ = \; & -\kappa_1 + \mu_1\lambda_-c_1^- - \mu_3\lambda_-c_3^-\left(\frac{\mu_1\lambda_-c_1^-}{\kappa_1 - \mu_1\lambda_-c_1^-}\right) < 0. \end{aligned} \tag{69}$$

Therefore, the conditions for the covariance stationarity are

$$0 < \kappa_2 < \frac{2}{\Delta t}, \tag{70}$$

$$\mu_1\lambda_-c_1^- < \kappa_1, \tag{71}$$

$$\frac{\mu_3\lambda_-c_3^-\kappa_1}{\kappa_1 - \mu_1\lambda_-c_1^-} < \kappa_3, \tag{72}$$

where

$$\lambda_- = \int_{-\infty}^{0} x^2\lambda_-e^{-\lambda_-|x|}dx = \frac{2}{\lambda_-^2}.$$

Interestingly, these stationarity conditions boil down to the following inequalities:

$$0 < \kappa_2, \tag{73}$$

$$\frac{2\mu_1c_1^-}{\lambda_-^2} < \kappa_1, \tag{74}$$

$$\frac{2\mu_3c_3^-\kappa_1}{\kappa_1\lambda_-^2 - 2\mu_1c_1^-} < \kappa_3, \tag{75}$$

which matches exactly the restrictions derived by Andersen et al. (2015a) in their footnote 37 (except the typo for κ_2).

Positivity We impose the classic feller condition

$$\sigma_1^2 \leq 2\kappa_1v_1, \quad \sigma_2^2 \leq 2\kappa_2v_2. \tag{76}$$

It is convenient to ensure that all element of $V_{0|0} = -K_1^{-1}\Phi_0$ are positive, where

$$\Phi_0 \equiv \Delta t \begin{pmatrix} \kappa_1 v_1 + \mu_1 \lambda_- c_0^- \\ \kappa_2 v_2 \\ \mu_3 \lambda_- c_0^- \end{pmatrix}, \quad K_1 = \Delta t \begin{bmatrix} -\kappa_1 + \mu_1 \lambda_- c_1^- & \mu_1 \lambda_- c_2^- & \mu_1 \lambda_- c_3^- \\ 0 & -\kappa_2 & 0 \\ \mu_3 \lambda_- c_1^- & \mu_3 \lambda_- c_2^- & -\kappa_3 + \mu_3 \lambda_- c_3^- \end{bmatrix}.$$

Because $\Phi_0 > 0$, the sign of $V_{0|0}$ is determined by the sign of $-K_1^{-1}$. We observe that

$$\begin{aligned} \det(K_1) &= -\kappa_2 \left(-\kappa_1 + \mu_1 \lambda_- c_1^-\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3^-\right) + \kappa_2 \mu_3 \lambda_- c_1^- \mu_1 \lambda_- c_3^-, \\ &= \kappa_2 \left(\mu_3 \lambda_- c_1^- \mu_1 \lambda_- c_3^- - \left(\kappa_1 - \mu_1 \lambda_- c_1^-\right)\left(\kappa_3 - \mu_3 \lambda_- c_3^-\right)\right), \\ &= \kappa_2 \left(-\kappa_3 \left(\kappa_1 - \mu_1 \lambda_- c_1^-\right) + \kappa_1 \mu_3 \lambda_- c_3^-\right), \\ &= -\kappa_2 \left(\kappa_1 - \mu_1 \lambda_- c_1^-\right)\left(\kappa_3 - \frac{\kappa_1 \mu_3 \lambda_- c_3^-}{\left(\kappa_1 - \mu_1 \lambda_- c_1^-\right)}\right) < 0, \end{aligned} \tag{77}$$

given the stationarity conditions. It follows that

$$K_1^{-1} = \frac{1}{\det(K_1)} \begin{bmatrix} -\kappa_2 \left(-\kappa_3 + \mu_3 \lambda_- c_3\right) & \mu_1 \lambda_- c_3 \mu_3 \lambda_- c_2 - \mu_1 \lambda_- c_2 \left(-\kappa_3 + \mu_3 \lambda_- c_3\right) & \kappa_2 \mu_1 \lambda_- c_3 \\ 0 & \left(-\kappa_1 + \mu_1 \lambda_- c_1\right)\left(-\kappa_3 + \mu_3 \lambda_- c_3\right) - \mu_1 \lambda_- c_3 \mu_3 \lambda_- c_1 & 0 \\ \kappa_2 \mu_3 \lambda_- c_1 & \mu_3 \lambda_- c_1 \mu_1 \lambda_- c_2 - \mu_3 \lambda_- c_2 \left(-\kappa_1 + \mu_1 \lambda_- c_1\right) & -\kappa_2 \left(-\kappa_1 + \mu_1 \lambda_- c_1\right) \end{bmatrix},$$

which simplifies to

$$K_1^{-1} = \frac{1}{\det(K_1)} \begin{bmatrix} \kappa_2 \left(\kappa_3 - \mu_3 \lambda_- c_3\right) & \mu_1 \lambda_- c_2 \kappa_3 & \kappa_2 \mu_1 \lambda_- c_3 \\ 0 & \kappa_3 \left(\kappa_1 - \mu_1 \lambda_- c_1\right) - \kappa_1 \mu_3 \lambda_- c_3 & 0 \\ \kappa_2 \mu_3 \lambda_- c_1 & \kappa_1 \mu_3 \lambda_- c_2 & \kappa_2 \left(\kappa_1 - \mu_1 \lambda_- c_1\right) \end{bmatrix}. \tag{78}$$

Given that

$$\kappa_3 > \frac{\kappa_1 \mu_3 \lambda_- c_3^-}{\left(\kappa_1 - \mu_1 \lambda_- c_1^-\right)} > \frac{\kappa_1 \mu_3 \lambda_- c_3^-}{\kappa_1} = \mu_3 \lambda_- c_3^-, \tag{79}$$

we conclude that all the element of K_1^{-1} are negative, and therefore, $V_{0|0}$ is always positive.

Table 1: **Descriptive statistics for risk-neutral moments**

This table reports descriptive statistics for risk-neutral volatility (Panel A), skewness (Panel B), and kurtosis (Panel C). Values are expressed in annualized percentage whenever appropriate. The maturity is denoted by τ expressed in months.

	τ 1	2	3	6	9	12	18	24
Panel A: Volatility								
Mean (%)	17.6814	18.8382	19.3424	19.9914	20.2686	20.3300	20.5658	20.9558
Std. Dev. (%)	7.5245	7.3475	7.1219	6.6033	6.3637	6.1460	6.0002	6.1472
AR(1)	0.9593	0.9777	0.9830	0.9867	0.9889	0.9862	0.9779	0.9593
Panel B: Skewness								
Mean	-1.7943	-1.8596	-1.7729	-1.7245	-1.6876	-1.6272	-1.6234	-1.6812
Std. Dev.	0.4936	0.4656	0.4490	0.3723	0.3529	0.3689	0.4320	0.5162
AR(1)	0.6986	0.9024	0.9065	0.8471	0.8270	0.7800	0.7582	0.6365
Panel C: Kurtosis								
Mean	7.4063	8.0706	7.6644	7.2955	6.8489	6.3382	6.1461	6.2482
Std. Dev.	2.8463	2.4007	2.1553	1.9084	1.8533	1.8372	2.1256	2.3903
AR(1)	0.6755	0.8225	0.7890	0.7387	0.7376	0.6888	0.6864	0.5452

Table 2: **Parameter estimates for the AFT model**

This table reports the parameter estimates with the corresponding standard errors (Std. Err.) for the AFT model, using risk-neutral moment-based and Andersen et al. (2015a) estimation approaches. The parameter values for $c_{\mathbb{Q}}$, $c_{\mathbb{Q}}^{\dagger}$, and η are set to 0, as in Andersen et al. (2015a).

	Moment-based		AFT Estimation	
Parameter	Estimate	Std. Err.	Estimate	Std. Err.
ρ_1	-0.9730	0.3485	-0.9860	0.0027
v_1	0.0178	0.0114	0.0145	0.0055
κ_1	1.2113	0.0155	1.9803	0.0577
σ_1	0.1607	0.0350	0.2350	0.0145
ρ_2	-0.6869	1.8001E-06	-0.8686	1.4430E-06
v_2	0.0640	0.0119	0.0285	0.0034
κ_2	1.1045	0.0098	1.1382	0.0507
σ_2	0.3558	0.0392	0.2436	0.0161
ρ_3	1.1059	0.5687	0.5934	0.2693
κ_3	0.2633	0.0204	0.2244	0.0138
ρ_3	0.0005	0.0001	0.0004	1.1527E-05
$c_{\mathbb{Q}}^{\dagger}$	2.3511	0.6336	1.7923	0.6575
c_1	10.0826	1.2052	22.8678	5.4523
$c_1^{\dagger}$	106.2423	22.2951	173.1724	53.4154
c_2	1.4435	0.2630	1.2769	0.1539
$c_2^{\dagger}$	72.2156	0.0005	56.3836	0.0004
c_3	66.4210	15.7729	72.0413	20.8272
λ	34.8709	9.4756	38.4586	7.8451
λ_1	132.7487	21.0389	156.7260	45.3644
μ_1	7.4634	4.7926	11.7914	4.4047

Table 3: **Vega-weighted root mean square errors for the AFT model-implied option prices**

This table presents vega-weighted root mean square errors (VWRMSE in percentages) for the AFT model-implied option prices sorted by moneyness and maturity. The moneyness denoted X/S is measured by the ratio of the strike price ($X = e^{r}$) to the underlying asset value (S). DTM denotes the number of calendar days to maturity and we include options with maturities of 1 month to 2 years. VWRMSEs are present for both risk-neutral moment-based (1) and Andersen et al. (2015a) (2) estimation approaches. The ratio of VWRMSEs between these two estimation approaches is also given below. A ratio (2)/(1) greater than 1 indicates that the risk-neutral moment-based estimation approach outperforms the Andersen et al. (2015a) estimation method in matching the observed option values.

Panel A: VWRMSE By Moneyness

Estimation	X/S<0.97	0.97<X/S<0.99	0.99<X/S<1.01	1.01<X/S<1.03	1.03<X/S<1.05	X/S>1.05	All
Moment-based (1)	5.2133	0.9555	0.9000	0.7625	0.5995	2.6720	3.6469
AFT estimation (2)	4.0184	0.9289	0.8484	0.7001	0.6803	3.5266	3.0868
Ratio (2)/(1)	0.7708	0.9722	0.9426	0.9182	1.1348	1.3198	0.8464

Panel B: VWRMSE By Maturity

	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<150	DTM>150	All
Moment-based (1)	1.4991	2.0397	2.3130	2.6891	3.1781	6.6370	3.6469
AFT estimation (2)	1.4206	1.9032	2.1382	2.6365	3.1948	5.3470	3.0868
Ratio (2)/(1)	0.9476	0.9331	0.9244	0.9804	1.0052	0.8056	0.8464

Table 4: **Shifts in Drift Parameter estimates for the AFT model**

This table reports the drift parameter estimates with the corresponding standard errors (Std. Err.) for the AFT model, under risk-neutral (Q) and physical (P) measures. The parameters are obtained from the risk-neutral moment-based estimation. Note that common structure preserving transformations restrict other model parameters such as σ_1, σ_2, η, ρ_1, ρ_2, ρ_3, μ_1, and κ_3 to be identical under both physical and risk-neutral probability measures.

Risk-Neutral (Q)			Physical (P)		
Parameter	Estimate	Std. Err.	Parameter	Estimate	Std. Err.
v_1^Q	0.0178	0.0114	v_1^P	0.0309	0.0086
κ_1^Q	1.2113	0.0155	κ_1^P	0.9713	0.0195
v_2^Q	0.0640	0.0119	v_2^P	0.0323	0.0215
κ_2^Q	1.1045	0.0098	κ_2^P	2.0624	1.3976

Figure 1: Factor representation of cumulants and leverage in the Heston (1993) model

These graphs plot the location $A_\tau^{(n)}$ and slope $B_\tau^{(n)}$ coefficients in the factor representation of cumulants ($CUM_\tau^{(n)}, n = 2, 3, 4$) against different values of leverage ($\rho \in [-1, 1]$) in the Heston (1993) model. The other model parameter values are $a = 0.02$, $b = 2$, $\sigma = 0.1$, and $\tau = 0.5$.

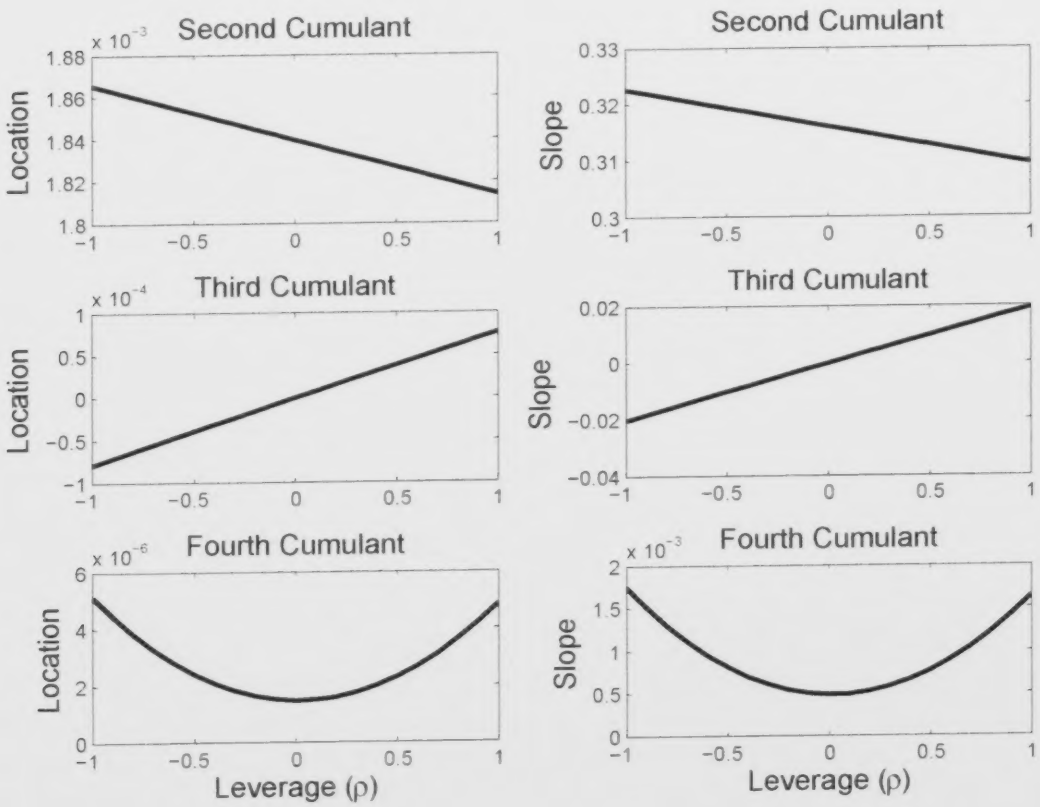

Figure 2: Factor representation of cumulants and first volatility factor leverage in the AFT model

These graphs plot the location $A_\tau^{(n)}$ and slope $B_\tau^{(n)}$ coefficients in the factor representation of cumulants ($CUM_\tau^{(n)}, n = 2, 3, 4$) against different values of leverage from the first volatility factor ($\rho_1 \in [-1, 1]$) in the Andersen et al. (2015a) estimation framework. The other model parameter values are $c_1 = 0.0084$, $\kappa_1 = 9.7196$, $\sigma_1 = 0.3924$, $\mu_1 = 13.4143$, $\rho_2 = -0.8707$, $c_2 = 0.0391$, $\kappa_2 = 0.1680$, $\sigma_2 = 0.1078$, $\mu_3 = 0.9238$, $\kappa_3 = 0.5967$, $\rho_3 = 0.0005$, $c_0^- = 0$, $c_0^+ = 1.5713$, $c_1^- = 25.3536$, $c_1^+ = 92.4094$, $c_2^- = 0.8802$, $c_2^+ = 72.5628$, $c_3^- = 41.4017$, $c_3^+ = 0$, $\lambda_- = 18.7455$, $\lambda_+ = 58.2399$, $\mu_1 = 13.4143$, $\eta = 0$, and $\tau = 0.5$.

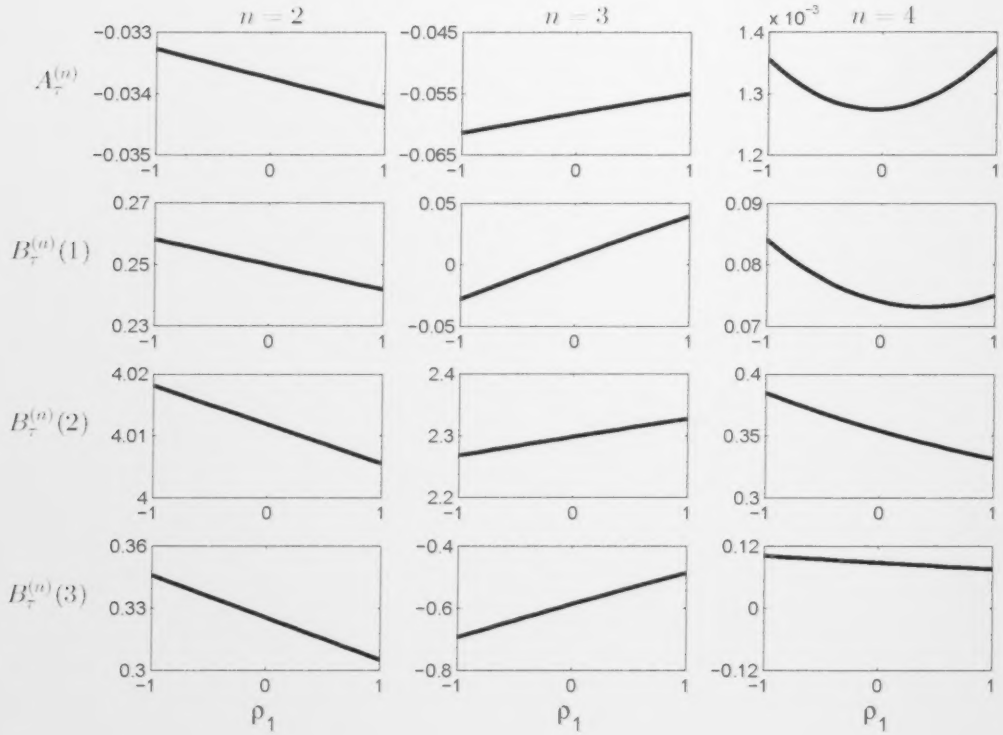

Figure 3: Factor representation of cumulants and second volatility factor leverage in the AFT model

These graphs plot the location $A_\tau^{(n)}$ and slope $B_\tau^{(n)}$ coefficients in the factor representation of cumulants ($CUM_\tau^{(n)}, n = 2, 3, 4$) against different values of leverage from the second volatility factor ($\rho_2 \in [-1, 1]$) in the Andersen et al. (2015a) estimation framework. The other model parameter values are $\rho_1 = -0.9818$, $\bar{v}_1 = 0.0084$, $\kappa_1 = 9.7196$, $\sigma_1 = 0.3924$, $\bar{v}_2 = 0.0391$, $\kappa_2 = 0.1680$, $\sigma_2 = 0.1078$, $\mu_3 = 0.9238$, $\kappa_3 = 0.5967$, $\rho_3 = 0.0005$, $c_0^- = 0$, $c_0^+ = 1.5713$, $c_1^- = 25.3536$, $c_1^+ = 92.4094$, $c_2^- = 0.8802$, $c_2^+ = 72.5628$, $c_3^- = 41.4017$, $c_3^+ = 0$, $\lambda_- = 18.7455$, $\lambda_+ = 58.2399$, $\mu_1 = 13.4143$, $\eta = 0$, and $\tau = 0.5$.

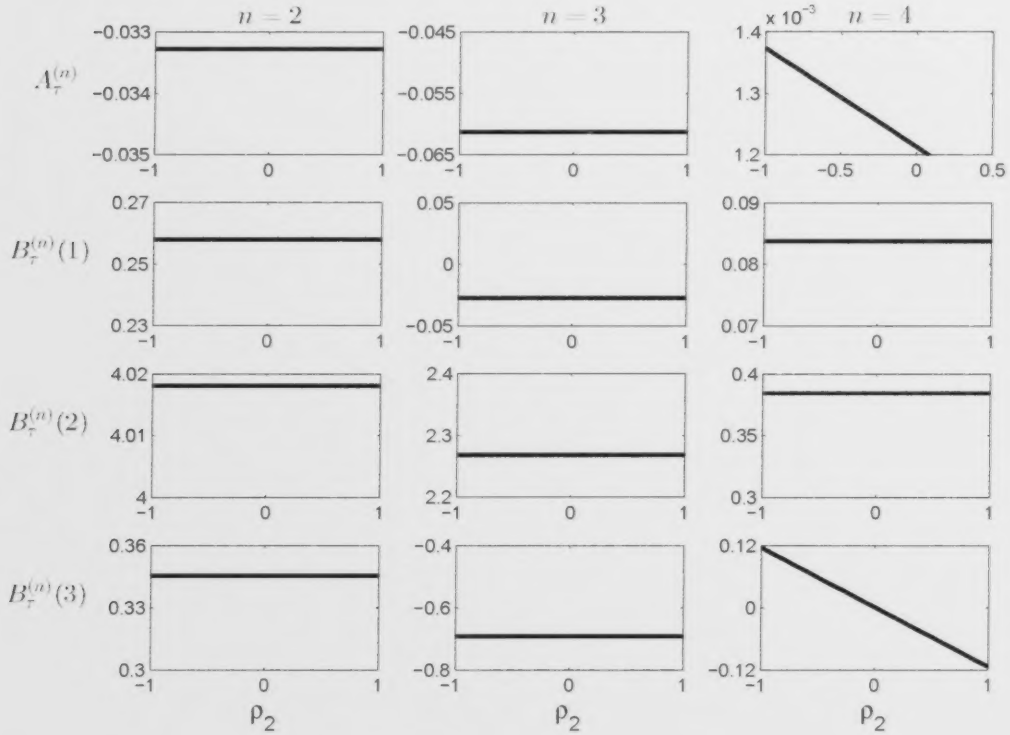

Figure 4: Observed risk-neutral moment dynamics

This graph displays the time series of observed risk-neutral volatility (left column), skewness (middle column), and kurtosis (right column) for 1-month (top row), 6-month (middle row), and 12-month (bottom row) maturities from September 03, 1996, to December 30, 2011.

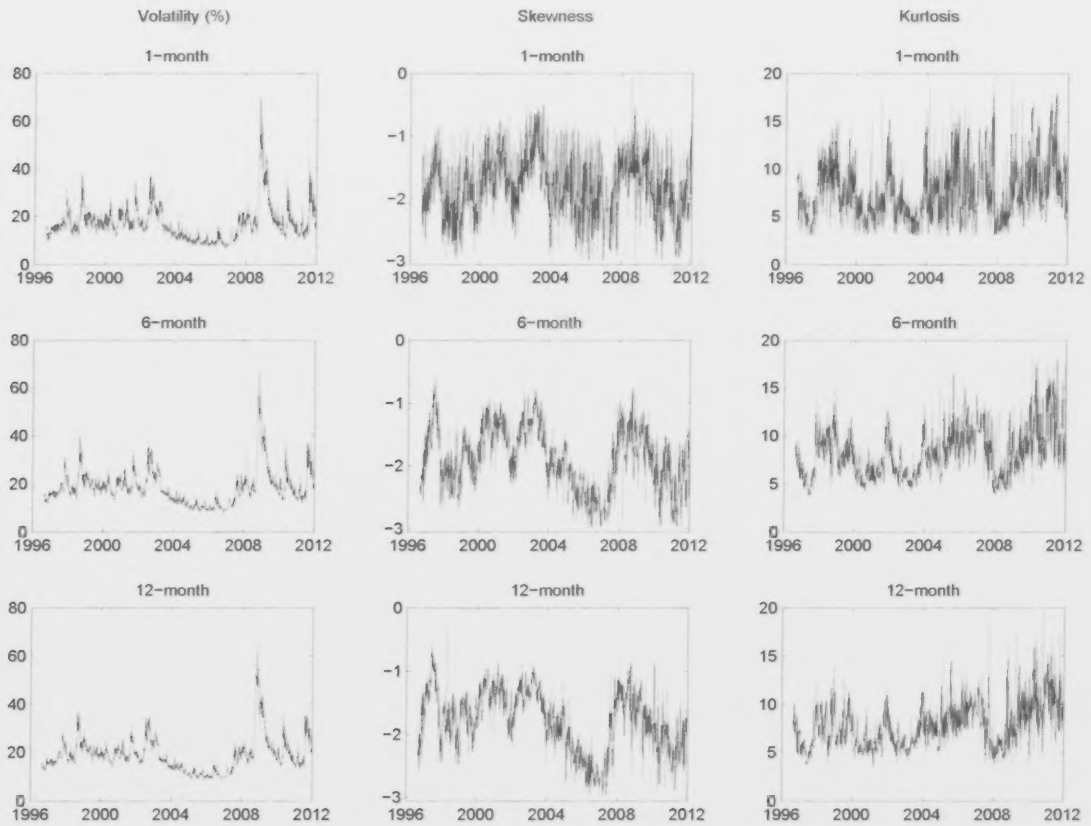

Figure 5: Model-implied factors

This graph shows the time series of AFT model-implied factors from September 03, 1996, to December 30, 2011. The factors are filtered using the Andersen et al. (2015a) (solid lines) and the moment-based (dotted lines) estimation approaches.

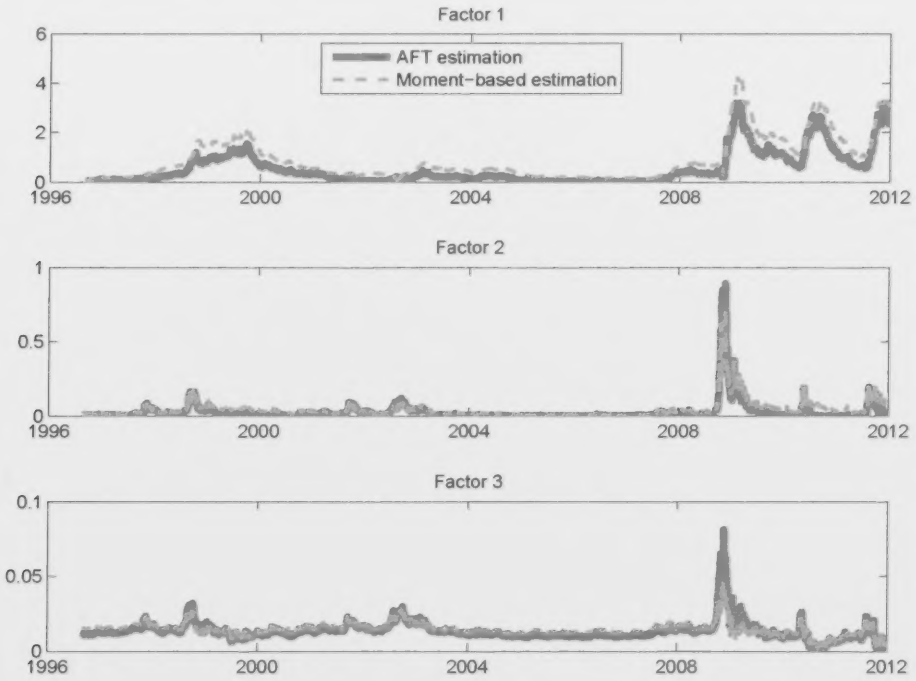

Figure 6: Model-implied risk-neutral moments

This graph plots the time series of AFT model-implied risk-neutral volatility (top), skewness (middle), and kurtosis (bottom) for 6-month maturity from September 03, 1996, to December 30, 2011. The observed series are constructed using the model-free procedure in Equation 8. We use the fitted parameter values in the second column of Table 2 to generate the implied series from the moment-based estimation approach. We use the fitted parameter values in the fourth column of Table 2 to generate the implied series from the Andersen et al. (2015a) estimation approach.

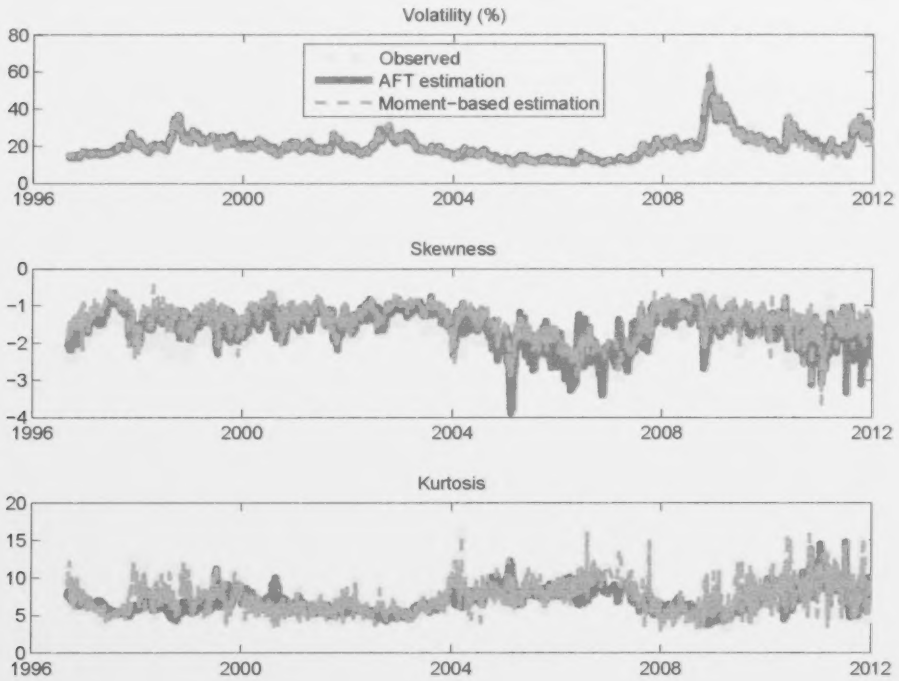

Figure 7: Model-implied moment risk premia

This graph displays the paths of the AFT model-implied risk-neutral (solid lines) and physical (dashed lines) variance (top), skewness (middle), and kurtosis (bottom) constructed from the moment-based estimation approach from September 03, 1996, to December 30, 2011. The shaded areas represent the wedge between the risk-neutral and physical series, and thus, reflect the corresponding premia.